



TRƯỜNG ĐẠI HỌC THƯƠNG MẠI

Chủ biên: TS. Đặng Văn Lương

GIÁO TRÌNH NGUYÊN LÝ THỐNG KÊ



NHÀ XUẤT BẢN THỐNG KÊ



TRƯỜNG ĐẠI HỌC THƯƠNG MẠI

Chủ biên: TS. Đặng Văn Lương

GIÁO TRÌNH NGUYÊN LÝ THỐNG KÊ



NHÀ XUẤT BẢN THỐNG KÊ

HÀ NỘI - 2016

LỜI MỞ ĐẦU

Với đường lối mở cửa và phát triển nền kinh tế thị trường định hướng xã hội chủ nghĩa, kinh tế Việt Nam trong những năm gần đây đã có những bước chuyển biến và phát triển. Điều đó đặt ra đòi hỏi ngày càng phải nâng cao chất lượng đào tạo trong các trường đại học kinh tế. Trong nền kinh tế thị trường cạnh tranh, công tác quản lý và ra các quyết định trong điều kiện không chắc chắn thì vai trò của thống kê ngày càng trở nên cần thiết và quan trọng. Để phù hợp với điều kiện thực tiễn, việc trang bị kiến thức thống kê cho các cử nhân kinh tế là yêu cầu không thể thiếu. Với tư cách là môn học cung cấp những kiến thức cơ bản về phương pháp thu thập, xử lý và phân tích thông tin phục vụ cho công tác quản lý, Nguyên lý thống kê là một học phần trong chương trình đào tạo của hầu hết của các trường đại học kinh tế nói chung và trường đại học thương mại nói riêng.

Để đáp ứng yêu cầu giảng dạy và học tập của giáo viên và sinh viên, bộ môn Thống kê - Phân tích tổ chức biên soạn giáo trình **Nguyên lý thống kê**. Đây là giáo trình phục vụ cho giảng dạy và học tập học phần Nguyên lý thống kê trong Trường Đại học Thương mại và làm tài liệu tham khảo cho các đối tượng quan tâm đến khoa học thống kê.

Giáo trình biên soạn lần này là sự kế thừa giáo trình Lý thuyết thống kê do bộ môn biên soạn trước đây và kết hợp những kinh nghiệm trong nhiều năm giảng dạy học phần này của bộ môn. Nội dung của giáo trình bao gồm những vấn đề lý luận và phương pháp thống kê cơ bản nhất, được trình bày dễ hiểu, kết hợp với những ví dụ minh họa cụ thể để phù hợp với đối tượng là sinh viên kinh tế ngoài ngành Thống kê.

Tham gia biên soạn gồm có:

TS. Đặng Văn Lương - Chủ biên - Biên soạn Chương 1 và chương 7

Ths. Đặng Thị Thu - Biên soạn Chương 5, chương 6

Ths. Phạm Thị Quỳnh Vân - Biên soạn Chương 4

CN. Nguyễn Văn Giao - Biên soạn Chương 3, Chương 8

Ths. Nguyễn Thị Mai - Biên soạn Chương 2

Mặc dù có nhiều cố gắng song không thể tránh khỏi những thiếu sót và hạn chế. Tập thể tác giả mong nhận được những ý kiến đóng góp để việc xuất bản giáo trình được hoàn thiện hơn.

Tập thể tác giả

MỤC LỤC

LỜI MỞ ĐẦU	3
CHƯƠNG 1: NHỮNG VẤN ĐỀ CHUNG VỀ THỐNG KÊ HỌC	11
1.1. Đối tượng nghiên cứu của thống kê học	11
1.1.1. Sơ lược sự ra đời và phát triển của thống kê học	12
1.1.2. Đối tượng nghiên cứu của thống kê học	13
1.2. Các khái niệm cơ bản của thống kê học	17
1.2.1. Tổng thể thống kê	17
1.2.2. Tiêu thức thống kê	19
1.2.3. Chỉ tiêu thống kê	21
1.2.4. Hệ thống chỉ tiêu thống kê	22
1.2.5. Dữ liệu thống kê	22
1.3. Thang đo trong thống kê	23
1.3.1. Thang đo định danh	23
1.3.2. Thang đo thứ bậc	23
1.3.3. Thang đo khoảng	24
1.3.4. Thang đo tỷ lệ	24
1.4. Quá trình nghiên cứu thống kê	25
1.5. Tổ chức thống kê ở Việt Nam	26
1.5.1. Hệ thống tổ chức thống kê nhà nước ở Việt Nam	26
1.5.2. Chức năng, nhiệm vụ của cơ quan thống kê nhà nước Việt Nam	27
<i>Tóm tắt chương 1</i>	28
<i>Câu hỏi ôn tập chương 1</i>	29
CHƯƠNG 2: ĐIỀU TRA THỐNG KÊ	30
2.1. Khái niệm, ý nghĩa và yêu cầu của điều tra thống kê	30
2.1.1. Khái niệm điều tra thống kê	30

2.1.2. Ý nghĩa của điều tra thống kê	31
2.1.3. Những yêu cầu cơ bản của điều tra thống kê	32
2.2. Các loại điều tra thống kê	32
2.2.1. Điều tra thường xuyên và không thường xuyên	32
2.2.2. Điều tra toàn bộ và không toàn bộ	34
2.3. Phương pháp thu thập thông tin trong điều tra thống kê	36
2.3.1. Phương pháp đăng ký trực tiếp	36
2.3.2. Phương pháp phỏng vấn	36
2.4. Hình thức tổ chức điều tra thống kê	38
2.4.1. Báo cáo thống kê định kỳ	38
2.4.2. Điều tra chuyên môn	39
2.5. Xây dựng phương án điều tra	40
2.5.1. Xác định mục đích điều tra	41
2.5.2. Xác định đối tượng và đơn vị điều tra	42
2.5.3. Chọn thời điểm, thời kỳ và quyết định thời hạn điều tra	42
2.5.4. Xác định nội dung điều tra và thiết lập phiếu điều tra	43
2.5.5. Các danh mục và bảng phân loại	44
2.5.6. Loại điều tra và phương pháp thu thập thông tin	45
2.5.7. Lập kế hoạch tổ chức và tiến hành điều tra	46
2.6. Xây dựng bảng hỏi trong điều tra thống kê	46
2.6.1. Bảng hỏi và yêu cầu của việc xây dựng bảng hỏi trong điều tra thống kê	46
2.6.2. Các loại câu hỏi và kỹ thuật đặt các loại câu hỏi	47
2.7. Sai số trong điều tra thống kê	52
2.7.1. Khái niệm	52
2.7.2. Biện pháp hạn chế sai số	53
Tóm tắt chương 2	54
Câu hỏi ôn tập chương 2	55
Bài tập chương 2	55

(CHƯƠNG 3: TỔNG HỢP THỐNG KÊ	58
33.1. Những vấn đề chung của tổng hợp thống kê	58
3.1.1. Khái niệm, ý nghĩa và nhiệm vụ của tổng hợp thống kê	58
3.1.2. Các vấn đề chủ yếu của tổng hợp thống kê	59
33.2. Phân tổ thống kê	62
3.2.1. Khái niệm, ý nghĩa, nhiệm vụ của phân tổ thống kê	62
3.2.2. Các bước tiến hành phân tổ thống kê	64
3.2.3. Dây số phân phối	71
33.3. Bảng thống kê và đồ thị thống kê	73
3.3.1. Bảng thống kê	73
3.3.2. Đồ thị thống kê	78
<i>Tóm tắt chương 3</i>	83
<i>Câu hỏi ôn tập chương 3</i>	84
<i>Bài tập chương 3</i>	84
 (CHƯƠNG 4: THỐNG KÊ MỨC ĐỘ CỦA HIỆN TƯỢNG KINH TẾ - XÃ HỘI	 87
44.1. Số tuyệt đối trong thống kê	87
4.1.1. Khái niệm và ý nghĩa của số tuyệt đối	87
4.1.2. Các loại số tuyệt đối	89
44.2. Số tương đối	89
4.2.1. Khái niệm và ý nghĩa của số tương đối	89
4.2.2. Các loại số tương đối	90
4.2.3. Điều kiện vận dụng số tuyệt đối và số tương đối	96
44.3. Số trung bình trong thống kê	96
4.3.1. Khái niệm, ý nghĩa của số trung bình	96
4.3.2. Các loại số trung bình	98
4.3.3. Điều kiện vận dụng số trung bình	110

4.4. Nghiên cứu độ biến thiên của tiêu thức	111
4.4.1. Ý nghĩa nghiên cứu độ biến thiên của tiêu thức	111
4.4.2. Các chỉ tiêu đo độ biến thiên của tiêu thức	111
<i>Tóm tắt chương 4</i>	<i>115</i>
<i>Câu hỏi ôn tập chương 4</i>	<i>116</i>
<i>Bài tập chương 4</i>	<i>116</i>
CHƯƠNG 5: HỒI QUY VÀ TƯƠNG QUAN	121
5.1. Mối liên hệ giữa các hiện tượng, nhiệm vụ của phương pháp hồi quy và tương quan	121
5.1.1. Mối liên hệ giữa các hiện tượng	121
5.1.2. Nhiệm vụ của phương pháp hồi quy và tương quan	123
5.2. Liên hệ tương quan tuyến tính giữa hai tiêu thức số lượng	125
5.2.1. Phương trình hồi quy tuyến tính	125
5.2.2. Hệ số tương quan	128
5.3. Liên hệ tương quan phi tuyến tính giữa hai tiêu thức số lượng	131
5.3.1. Các phương trình hồi quy phi tuyến tính	131
5.3.2. Tỷ số tương quan	134
5.4. Liên hệ tương quan tuyến tính giữa nhiều tiêu thức số lượng	138
5.4.1. Phương trình hồi quy tuyến tính nhiều tiêu thức	139
5.4.2. Hệ số tương quan	140
5.5. Hệ số co giãn	144
<i>Tóm tắt chương 5</i>	<i>145</i>
<i>Câu hỏi ôn tập chương 5</i>	<i>146</i>
<i>Bài tập chương 5</i>	<i>147</i>
CHƯƠNG 6: DÃY SỐ THỜI GIAN	150
6.1. Khái niệm, ý nghĩa của dãy số thời gian	150
6.2. Các chỉ tiêu phân tích dãy số thời gian	153
6.2.1. Mức độ trung bình theo thời gian	153

6.2.2. Lượng tăng (giảm) tuyệt đối	155
6.2.3. Tốc độ phát triển	158
6.2.4. Tốc độ tăng (giảm)	160
6.2.5. Giá trị tuyệt đối 1% tăng (giảm)	162
6.3. Các phương pháp biểu hiện xu hướng phát triển cơ bản của hiện tượng	163
6.3.1. Phương pháp mở rộng khoảng cách thời gian	164
6.3.2. Phương pháp số trung bình di động (số trung bình trượt)	165
6.3.3. Phương pháp hồi quy	167
6.3.4. Phương pháp biểu hiện biến động thời vụ	171
6.4. Dự báo thống kê	173
6.4.1. Khái niệm và phân loại dự báo	173
6.4.2. Dự báo thống kê	175
6.4.3. Một số phương pháp dự báo thống kê thông dụng	176
Tóm tắt chương 6	178
Câu hỏi ôn tập chương 6	179
Bài tập chương 6	180
 CHƯƠNG 7: CHỈ SỐ	 184
7.1. Một số vấn đề chung về chỉ số	184
7.1.1. Khái niệm, đặc điểm và ý nghĩa của chỉ số	184
7.1.2. Phân loại chỉ số	186
7.2. Phương pháp tính chỉ số	188
7.2.1. Phương pháp tính chỉ số phát triển	188
7.2.2. Phương pháp tính chỉ số không gian	199
7.3. Hệ thống chỉ số	200
7.3.1. Hệ thống chỉ số tổng hợp	201
7.3.2. Hệ thống chỉ số phân tích biến động của chỉ tiêu trung bình	205
7.3.3. Hệ thống chỉ số nghiên cứu biến động của chỉ tiêu tổng lượng biến	209

7.4. Một số chỉ số thông dụng ở Việt Nam	2122
7.4.1. Chỉ số giá tiêu dùng	2122
7.4.2. Chỉ số giá chứng khoán	2133
7.4.3. Chỉ số sản xuất công nghiệp	2144
<i>Tóm tắt chương 7</i>	2166
<i>Câu hỏi ôn tập chương 7</i>	2177
<i>Bài tập chương 7</i>	2188
 CHƯƠNG 8: ĐIỀU TRA CHỌN MẪU	2222
8.1. Khái niệm và ý nghĩa của điều tra chọn mẫu	2222
8.2. Điều tra chọn mẫu ngẫu nhiên	2255
8.2.1. Một số lý luận trong điều tra chọn mẫu ngẫu nhiên	2255
8.2.2. Các phương pháp tổ chức chọn mẫu ngẫu nhiên thường dùng trong thống kê	2383
8.3. Quy trình một cuộc điều tra chọn mẫu ngẫu nhiên	2455
8.3.1. Xác định mục đích nghiên cứu	2455
8.3.2. Xác định tổng thể nghiên cứu	2455
8.3.3. Xác định nội dung điều tra	2455
8.3.4. Xác định số lượng đơn vị của tổng thể mẫu và phương pháp tổ chức chọn mẫu	2466
8.3.5. Tiến hành thu thập tài liệu ở các đơn vị của tổng thể mẫu	2466
8.3.6. Suy rộng kết quả điều tra chọn mẫu	2466
8.3.7. Đưa ra kết luận về tổng thể chung	2477
8.4. Điều tra chọn mẫu phi ngẫu nhiên	2477
8.4.1. Phân tổ chính xác hiện tượng nghiên cứu	2477
8.4.2. Xác định số lượng đơn vị cần điều tra	2477
8.4.3. Lựa chọn các đơn vị điều tra	2488
8.4.4. Suy rộng kết quả điều tra	2488
<i>Tóm tắt chương 8</i>	2488
<i>Câu hỏi ôn tập chương 8</i>	2499
<i>Bài tập chương 8</i>	2499

Chương 1

INHỮNG VẤN ĐỀ CHUNG VỀ THỐNG KÊ HỌC

Thống kê là khoa học thu thập, tổng hợp, phân tích và dự đoán dữ liệu. Thông tin thống kê có vai trò quan trọng phục vụ cho công tác quản lý vĩ mô cũng như trong doanh nghiệp. Là một môn khoa học độc lập, thống kê có đối tượng và phương pháp nghiên cứu riêng. Trong chương này sẽ đề cập những vấn đề khái quát chung của khoa học thống kê, bao gồm: Sơ lược lịch sử ra đời và phát triển, đối tượng nghiên cứu, một số khái niệm cơ bản và tổ chức công tác thống kê hiện nay ở Việt Nam. Nghiên cứu nội dung chương này giúp người học có được nhận thức về những vấn đề tổng quan chung của khoa học thống kê.

1.1. Đối tượng nghiên cứu của thống kê học

Thuật ngữ thống kê được hiểu với nhiều nghĩa khác nhau:

Thống kê là một môn khoa học, nghiên cứu việc biểu hiện các hiện tượng kinh tế, xã hội bằng con số. Thống kê là một công cụ của nhận thức, được sử dụng trong lĩnh vực tự nhiên, xã hội.

Thống kê là một hình thức hoạt động thực tiễn của con người nhằm thu thập, tổng hợp, phân tích các số liệu thống kê về dân số, kinh tế, văn hoá, giáo dục và các lĩnh vực khác của đời sống xã hội.

Giữa khoa học thống kê và công tác thống kê có mối liên hệ mật thiết với nhau. Khoa học thống kê sử dụng số liệu thực tiễn để củng cố về phương pháp và phương pháp luận. Ngược lại, thống kê thực tiễn sử dụng khoa học thống kê để giải quyết các yêu cầu thực tế của đời sống xã hội.

Trong lịch sử, những ý tưởng và phương pháp thống kê phát triển cùng với sự phát triển trên cơ sở nhu cầu của xã hội về thu thập và sử dụng thông tin. Do vậy, trước khi nghiên cứu nội dung môn học, việc tìm hiểu sơ lược về sự ra đời và phát triển của khoa học thống kê là cần thiết.

1.1.1. Sơ lược sự ra đời và phát triển của thống kê học

Thống kê học là môn khoa học ra đời và phát triển do nhu cầu thực tiễn của xã hội. Để trở thành một môn khoa học phát triển như ngày nay, nó đã phải trải qua một quá trình phát triển lâu dài gắn liền với những mốc lịch sử phát triển của xã hội loài người. Đây là một quá trình đúc kết dần kinh nghiệm thành lý luận khoa học hoàn thiện như ngày nay.

Mỗi một mốc lịch sử phát triển xã hội có ảnh hưởng rất lớn tới hoạt động thống kê, vì Thống kê học phục vụ cho lợi ích của một giai cấp nhất định.

Ngay từ thời cổ đại, con người đã biết chú ý tới việc đăng ký, ghi chép và tính toán số người trong bộ tộc, số hoa quả hái lượm được, số súc vật săn bắt được... Mặc dù việc ghi chép còn rất đơn giản và thô sơ trên một phạm vi rất hẹp, nhưng nó chính là những hoạt động ban đầu mang tính chất thống kê.

Trong xã hội phong kiến, hầu hết các quốc gia Á, Âu đều tổ chức việc đăng ký, kê khai số dân, ruộng đất, tài sản... với phạm vi rộng hơn, có tính thống kê rõ hơn. Tuy nhiên, công tác đăng ký này còn mang tính tự phát, thiếu khoa học. Ở giai đoạn này thống kê đã có những bước phát triển rõ rệt nhưng chưa thực sự trở thành một môn khoa học độc lập.

Sự ra đời và phát triển mạnh mẽ của phương thức sản xuất tư bản chủ nghĩa, đòi hỏi các nhà khoa học phải nghiên cứu lý luận cũng như phương pháp thu thập, tính toán và phân tích mặt lượng của các hiện tượng kinh tế - xã hội. Năm 1660, nhà kinh tế học người Đức H. Conring đã giảng về phương pháp nghiên cứu hiện tượng xã hội dựa vào số liệu điều tra cụ thể. Năm 1682, William Petty, nhà kinh tế học người Anh đã xuất bản cuốn “Số học chính trị”; đây là tác phẩm thống kê đầu tiên, trong đó tác giả nghiên cứu các hiện tượng xã hội bằng cách tổng hợp, so sánh các con số. *Karl Marx* đã gọi *William Petty* là người sáng lập ra môn thống kê học.

Trong những giai đoạn lịch sử tiếp theo, với sự ra đời của “Lý thuyết xác suất và thống kê toán” đã có sự ảnh hưởng lớn đến sự phát triển của thống kê học. Kể từ đó thống kê có sự phát triển mạnh mẽ và ngày càng hoàn thiện về lý luận và phương pháp.

Ngày nay, thống kê được coi là một trong những công cụ quản lý vĩ mô quan trọng, có vai trò cung cấp thông tin thống kê trung thực, khách quan, chính xác, đầy đủ, kịp thời phục vụ các cơ quan nhà nước trong việc đánh giá, dự báo tình hình, hoạch định chiến lược, chính sách, xây dựng kế hoạch phát triển kinh tế - xã hội ngắn hạn và dài hạn.

Trên giác độ vi mô, thống kê có vai trò cung cấp thông tin đánh giá mặt lượng các hoạt động kinh tế - xã hội của các đơn vị cơ sở. Những thông tin thống kê có vai trò quan trọng phục vụ cho công tác quản lý và ra quyết định đối với mọi cấp, mọi ngành cũng như trong tổ chức hoạt động sản xuất - kinh doanh của doanh nghiệp.

1.1.2. Đối tượng nghiên cứu của thống kê học

Mỗi môn khoa học độc lập đều có đối tượng nghiên cứu riêng của nó, đối tượng nghiên cứu phản ánh khái quát nội dung nghiên cứu của môn học và giới hạn phạm vi môn học. Sự giới hạn về phạm vi nghiên cứu giúp hình thành nên những ranh giới để phân biệt môn học này với môn học khác. Nắm vững được đối tượng nghiên cứu của môn học, không những làm rõ nội dung cơ bản của môn học, thấy được nét đặc thù và qua đó hiểu rõ được ý nghĩa, tác dụng của môn học, đồng thời còn giúp xác định phương pháp nghiên cứu môn học.

Thống kê học là một môn khoa học có đối tượng và phương pháp nghiên cứu độc lập. Đối tượng nghiên cứu của khoa học thống kê được thể hiện qua những đặc trưng cơ bản sau đây:

+ *Thống kê học nghiên cứu các hiện tượng và quá trình kinh tế - xã hội. Các hiện tượng kinh tế - xã hội bao gồm:*

- *Về kinh tế:* Quá trình tái sản xuất mở rộng của cái vật chất xã hội, quá trình sản xuất của các ngành, tình hình sử dụng tài nguyên thiên nhiên, tình hình phân phối sản phẩm xã hội theo hình thức sở hữu, tình hình sử dụng tài nguyên và các sản phẩm xã hội cũng như hiệu quả của việc sử dụng nó, ...

- *Về xã hội:* Dân số (số nhân khẩu, cấu thành của nhân khẩu, tình hình biến động về nhân khẩu); Giáo dục: Trình độ văn hoá; Y tế: Tình hình chăm sóc sức khoẻ, tình hình tiêm chủng cho trẻ sơ sinh...

- *Về các hoạt động chính trị*: Số người tham gia biểu tình, mít tinh, bầu cử quốc hội, hội đồng nhân dân, cơ cấu đại biểu quốc hội, tình hình tội phạm, ...

Thống kê học chủ yếu nghiên cứu các hiện tượng kinh tế - xã hội, song hiện tượng tự nhiên có mối quan hệ chặt chẽ với các hiện tượng kinh tế - xã hội. Vì vậy, thống kê nghiên cứu ảnh hưởng của các yếu tố tự nhiên (điều kiện địa lý, thời tiết, khí hậu...) và của các yếu tố kỹ thuật (việc phát minh sáng chế, cải tiến công cụ sản xuất...) đối với sự phát triển sản xuất và điều kiện sinh hoạt xã hội.

Như vậy, phạm vi nghiên cứu của thống kê học rất rộng bao gồm cả những hiện tượng kinh tế - xã hội thuộc lực lượng sản xuất và quan hệ sản xuất, cả những hiện tượng kinh tế - xã hội thuộc cơ sở hạ tầng và kiến trúc thượng tầng.

+ *Thống kê học nghiên cứu mặt lượng, nhưng không phải mặt lượng đơn thuần mà là mặt lượng trong sự liên hệ mật thiết với mặt chất của hiện tượng kinh tế - xã hội.*

Không giống với các môn khoa học khác, thống kê học không trực tiếp nghiên cứu bản chất và quy luật của hiện tượng xã hội mà nghiên cứu nó thông qua những biểu hiện về lượng, cụ thể là con số thống kê.

Mỗi một hiện tượng bao giờ cũng có 2 mặt: Mặt chất và mặt lượng tồn tại trong quan hệ mật thiết không thể tách rời, giữa chúng có mối liên hệ biện chứng với nhau.

- *Mặt chất*: Đó là bản chất triu tượng của hiện tượng, được biểu hiện qua nội dung, ý nghĩa, tính quy luật, bản chất cụ thể. Chất của hiện tượng giúp ta phân biệt hiện tượng này với hiện tượng khác.

- *Mặt lượng*: Biểu hiện cụ thể của mặt chất được thể hiện bằng con số nói lên quy mô, kết cấu, quan hệ về lượng, tốc độ phát triển, trình độ phổ biến... của hiện tượng nghiên cứu. Mặt lượng cho ta thấy hiện tượng ở mức độ nào, trình độ phổ biến ra sao, lớn hay nhỏ, cao hay thấp, nặng hay nhẹ, nhanh hay chậm. Mỗi lượng cụ thể đều gắn với một chất nhất định, ngược lại một chất lại được biểu hiện bằng một lượng. Chính vì vậy, nghiên cứu lý luận cũng như trong hoạt động thực tiễn, việc tìm hiểu

cả hai mặt của hiện tượng đều quan trọng như nhau. Trong hoạt động thực tiễn nếu không hiểu đầy đủ mặt lượng của hiện tượng sẽ dẫn tới những quyết định xa rời thực tế.

Mặt chất của hiện tượng kinh tế - xã hội được nhiều môn khoa học nghiên cứu còn mặt lượng của nó là đối tượng nghiên cứu của thống kê học. Thông qua mặt lượng, thống kê giúp ta xác định bản chất cụ thể của hiện tượng nghiên cứu.

Các số liệu thống kê có thể phản ánh được mặt chất của hiện tượng, vì chất và lượng là hai mặt không thể tách rời nhau của sự vật hay hiện tượng, giữa chúng có mối liên hệ biện chứng với nhau. Quan điểm Triết học đã khẳng định: “Lượng là biểu hiện cụ thể của chất, chất tồn tại thông qua lượng, lượng đổi thì chất đổi”. Như vậy, hiện tượng không thể có chất mà không có lượng, ngược lại không thể có lượng mà không có chất, sự biến đổi về lượng dẫn đến sự thay đổi về chất. Chính vì vậy, nghiên cứu mặt lượng của hiện tượng có ý nghĩa to lớn đối với việc nhận thức bản chất và quy luật phát triển của hiện tượng.

Thống kê học nêu ra những lý luận và phương pháp để biểu hiện và phân tích được quy mô, kết cấu, quan hệ so sánh, trình độ phổ biến, tốc độ và xu hướng phát triển của hiện tượng kinh tế - xã hội. Nói một cách khác, thống kê biểu hiện mặt lượng của bản chất và quy luật phát triển của hiện tượng nghiên cứu. Ví dụ, để biết được quy mô kinh doanh của một doanh nghiệp thì phải trên cơ sở tính toán doanh nghiệp đó có bao nhiêu tài sản, có bao nhiêu công nhân, vốn cố định có bao nhiêu, vốn lưu động có bao nhiêu...

+ Hiện tượng kinh tế - xã hội mà thống kê nghiên cứu thường là hiện tượng số lớn.

Thống kê học coi tổng thể các hiện tượng cá biệt là một thể hoàn chỉnh, một tập hợp nhiều yếu tố kết hợp với nhau và lấy tổng thể hoàn chỉnh đó, cả tập hợp đó làm đối tượng nghiên cứu.

Mặt lượng của hiện tượng kinh tế - xã hội mà thống kê nghiên cứu không hoàn toàn giống nhau trên các đơn vị cấu thành hiện tượng, vì mỗi một đơn vị cá biệt có mặt lượng khác nhau. Lượng của hiện tượng cá biệt

chịu tác động của nhiều nhân tố, có những nhân tố bản chất, tất nhiên, cũng có những nhân tố không bản chất, ngẫu nhiên. Hướng tác động của các nhân tố này trên các hiện tượng cá biệt rất khác nhau, nhất là các nhân tố ngẫu nhiên (hướng tác động của các yếu tố ngẫu nhiên thường không cùng chiều với nhau). Vì vậy, chỉ có thông qua việc nghiên cứu một số lớn hiện tượng, tác động của các nhân tố ngẫu nhiên được bù trừ và triệt tiêu, biểu hiện bản chất và tính quy luật của hiện tượng kinh tế - xã hội mới có khả năng thể hiện rõ rệt. Trên cơ sở đó, mặt lượng sau khi tổng hợp có thể biểu hiện được bản chất và quy luật của hiện tượng nghiên cứu. Ví dụ: Khi nghiên cứu mức sống dân cư, mặc dù có thể thấy mức sống người dân, cùng với sự phát triển của kinh tế - xã hội, có xu hướng tăng lên, điều này có tính chất quy luật. Song, nếu chỉ nghiên cứu một vài gia đình thì không thể có được số liệu chính xác phản ánh tình hình đó. Có những gia đình mức sống tăng lên rất nhanh nhưng có những gia đình mức sống vẫn chưa được nâng cao. Vì vậy, bản chất và tính quy luật trên không thể hiện ra được, hoặc thể hiện không chính xác, không đảm bảo cơ sở khoa học. Nhưng nếu nghiên cứu một số lượng lớn gia đình thì tác động của các nhân tố không bản chất, ngẫu nhiên của từng gia đình cá biệt sẽ được bù trừ, triệt tiêu, bản chất và tính quy luật trên sẽ được biểu hiện bằng số lượng rõ ràng và có cơ sở khoa học.

Thống kê nghiên cứu hiện tượng số lớn không có nghĩa là không nghiên cứu hiện tượng cá biệt. Giữa hiện tượng số lớn và hiện tượng cá biệt tồn tại mối quan hệ biện chứng. Hơn nữa, trong quá trình phát triển của hiện tượng kinh tế thường phát sinh một số đơn vị (hiện tượng) cá biệt mới. Cho nên nghiên cứu hiện tượng số lớn kết hợp với nghiên cứu hiện tượng cá biệt là điều cần thiết. Nó giúp cho việc nhận thức hiện tượng được toàn diện, phong phú và sâu sắc. Đối với công tác quản lý kinh tế, xã hội việc nghiên cứu đơn vị và các cá nhân điển hình tiên tiến có ý nghĩa quan trọng.

+ Thống kê học nghiên cứu hiện tượng trong điều kiện thời gian và không gian cụ thể

Hiện tượng kinh tế, xã hội bao giờ cũng tồn tại trong những điều kiện thời gian và địa điểm (không gian) cụ thể. Trong những điều kiện

liệu sử khác nhau, hiện tượng có các đặc điểm về chất và biểu hiện về lượng khác nhau, điều này sẽ đảm bảo tính chính xác của số liệu thống kê. Tính cụ thể, chính xác của số liệu là mối quan tâm hàng đầu của thống kê.

Từ những đặc điểm cơ bản về đối tượng nghiên cứu của khoa học thống kê như đã trình bày ở trên, có thể khái quát: *Đối tượng nghiên cứu của thống kê học là mặt lượng trong sự liên hệ mật thiết với mặt chất của các hiện tượng kinh tế - xã hội số lớn, trong điều kiện thời gian và địa điểm cụ thể.*

Đối tượng của thống kê học quyết định đối tượng nghiên cứu của nguyên lý thống kê.

Nguyên lý thống kê (lý thuyết thống kê) là một môn khoa học nghiên cứu hệ thống các phương pháp thu thập, xử lý, phân tích các con số (mặt lượng) của những hiện tượng số lớn để tìm hiểu bản chất và tính quy luật vốn có của chúng (mặt chất) trong điều kiện thời gian và không gian cụ thể.

1.2. Các khái niệm cơ bản của thống kê học

Thống kê học dùng nhiều khái niệm chuyên môn để nghiên cứu đối tượng của mình. Trong nguyên lý thống kê thường gặp các khái niệm: Tổng thể, đơn vị tổng thể, tiêu thức, chỉ tiêu, hệ thống chỉ tiêu.

1.2.1. Tổng thể thống kê

1.2.1.1. Khái niệm tổng thể thống kê

Đối tượng nghiên cứu của thống kê thường là hiện tượng kinh tế - xã hội số lớn bao gồm nhiều đơn vị, nhiều hiện tượng cá biệt kết hợp với nhau. Thống kê học dùng khái niệm tổng thể thống kê để chỉ đối tượng nghiên cứu trong một trường hợp nghiên cứu cụ thể.

Tổng thể thống kê là hiện tượng số lớn, bao gồm những đơn vị hoặc phần tử cấu thành hiện tượng, cần được quan sát và phân tích mặt lượng của chúng.

Đây là khái niệm quan trọng trong thống kê, xác định tổng thể nhằm quy định về phạm vi nghiên cứu của hiện tượng. Từ đó có thể xác

định phạm vi điều tra, tổng hợp và phân tích số liệu về hiện tượng đó trong thời gian và địa điểm chính xác.

Ví dụ: Toàn bộ nhân khẩu nước ta là tổng thể, là tập hợp những công dân Việt Nam, không phân biệt tuổi tác, giới tính, tôn giáo, nghề nghiệp... mỗi người dân là một đơn vị tổng thể, gọi là tổng thể nhân khẩu Việt Nam.

Các doanh nghiệp thương mại cũng lập thành một tổng thể, chúng là tập hợp của những doanh nghiệp kinh doanh trong lĩnh vực thương mại mà không phân biệt hình thức sở hữu, quy mô, mặt hàng kinh doanh. Mỗi doanh nghiệp là 1 đơn vị tổng thể.

Như vậy, muốn xác định được một tổng thể thống kê, cần phải xác định được tất cả các đơn vị tổng thể thuộc phạm vi đó. Thực chất của việc xác định tổng thể thống kê là xác định các đơn vị tổng thể. Trong nhiều trường hợp thì đơn vị tổng thể là phần tử không thể chia nhỏ được. Đơn vị tổng thể là căn cứ quan trọng để xác định phương pháp điều tra thích hợp cũng như quá trình tổng hợp số liệu sau này.

1.2.1.2. Phân loại tổng thể

+ *Căn cứ vào hình thức biểu hiện có thể chia tổng thể thành 2 loại:*

- *Tổng thể bộc lộ:* Là tổng thể bao gồm các đơn vị mà ranh giới được biểu hiện rõ ràng, có thể xác định bằng trực quan. Ví dụ: Tổng thể nhân khẩu, tổng thể các sản phẩm bán ra của một công ty thương mại ở một thời điểm nào đó, tổng thể các trường đại học trong nước...

- *Tổng thể tiềm ẩn:* Là tổng thể mà ranh giới không rõ ràng và không thể nhận biết trực tiếp bằng trực quan. Ví dụ: Tổng thể những sinh viên có ý thức học tập tốt, những người có năng khiếu nghệ thuật, thể thao,... Loại tổng thể này thường gặp trong lĩnh vực xã hội.

Các đơn vị tổng thể thường chỉ giống nhau ở những đặc điểm cơ bản là đặc trưng của tổng thể, các đặc điểm còn lại có thể khác nhau nhiều hoặc ít.

+ *Căn cứ vào mục đích nghiên cứu chia ra:*

- *Tổng thể đồng chất:* Bao gồm các đơn vị giống nhau về đặc điểm chủ yếu có liên quan trực tiếp đến mục đích nghiên cứu.

- *Tổng thể không đồng chất*: Bao gồm các đơn vị có nhiều đặc điểm chủ yếu khác nhau.

Sự phân chia có ý nghĩa quan trọng trong việc xác định tính chất đại diện của các con số thống kê tính toán được, như trong trường hợp tính các chỉ tiêu trung bình.

+ *Căn cứ vào phạm vi có thể chia ra*:

- *Tổng thể chung*: Bao gồm tất cả các đơn vị của tổng thể.

- *Tổng thể bộ phận*: Chỉ bao gồm một phần của tổng thể chung.

Ví dụ: Nếu coi tổng thể các doanh nghiệp trên địa bàn thành phố Hà Nội là một tổng thể chung, thì tổng thể các doanh nghiệp trên chia theo ngành nghề kinh doanh hoặc theo hình thức sở hữu là các tổng thể bộ phận.

Trong thực tế, nhiều khi ranh giới của tổng thể không rõ ràng, do đó việc xác định tổng thể mang tính tương đối, quy ước không cố định. Để đáp ứng mục đích nghiên cứu, thống kê phải trên cơ sở phân tích lý luận kinh tế, chính trị, xã hội, định nghĩa rõ tổng thể. Định nghĩa tổng thể không những phải phản ánh giới hạn về thực thể (tổng thể đó là gì) mà còn phải giới hạn về thời gian (tổng thể tồn tại vào thời gian nào, ở đâu). Xác định tổng thể chính xác không dễ dàng vì có những hiện tượng có thể tương tự về hình thức, nhưng lại khác hẳn nhau về mặt nội dung. Chính vì vậy, phải phân tích lý luận để thấy rõ nội dung của hiện tượng. Xác định tổng thể không chính xác không những lãng phí sức người và tiền của mà còn không đủ cơ sở để hiểu đúng đắn bản chất cụ thể của vấn đề nghiên cứu.

1.2.2. Tiêu thức thống kê

1.2.2.1. Khái niệm

Nghiên cứu thống kê phải dựa vào các đặc điểm của đơn vị tổng thể. Mỗi đơn vị tổng thể có rất nhiều đặc điểm. Trong đó, có một số đặc điểm cấu thành tổng thể (tức là các đơn vị tổng thể đều có đặc điểm đó) và các đặc điểm khác. Ví dụ: Mỗi người dân trong tổng thể nhân khẩu nước ta có đặc điểm như: Họ, tên, tuổi, giới tính, trình độ văn hoá, trình độ chuyên môn, tiền lương,... Tùy theo mục đích nghiên cứu, một hay

một số đặc điểm được chọn ra làm cơ sở để thu thập dữ liệu, những đặc điểm đó gọi là tiêu thức.

Như vậy: Tiêu thức thống kê là một khái niệm chỉ đặc điểm của các đơn vị tổng thể được chọn ra để nghiên cứu.

1.2.2.2. Phân loại

Tiêu thức thống kê có thể chia thành:

- *Tiêu thức thuộc tính*: Là loại tiêu thức mà các biểu hiện của nó được dùng để phản ánh các thuộc tính (loại hình hoặc tính chất) của đơn vị tổng thể. Có nghĩa là không được biểu hiện trực tiếp bằng con số. Ví dụ: Trên một nhân khẩu có các tiêu thức thuộc tính như: Giới tính, tình trạng hôn nhân, nghề nghiệp, nơi cư trú, nhân cách...

Tiêu thức thuộc tính có thể biểu hiện trực tiếp, ví dụ tiêu thức nghề nghiệp (công chức, viên chức, công nhân...), tiêu thức giới tính (nam, nữ) hoặc biểu hiện gián tiếp. Ví dụ: Tiêu thức nhân cách, muốn đánh giá nhân cách mỗi người, ta phải đánh giá thông qua tiêu thức trung gian như mức độ trung thực, giúp đỡ mọi người... Tiêu thức mức sống dân cư được biểu hiện thông qua: GDP/người, sản lượng lương thực, thực phẩm, diện tích nhà ở/hộ...

- *Tiêu thức số lượng*: Là loại tiêu thức có biểu hiện trực tiếp bằng con số, đây là những con số phản ánh đặc trưng về số lượng của đơn vị tổng thể, có thể cân, đong, đo, đếm được và làm các phép tính như: Cộng, trừ, trung bình, tỷ lệ... Mỗi con số này được gọi là một lượng biến, các lượng biến chia ra lượng biến rời rạc và lượng biến liên tục. Ví dụ: Tiêu thức độ tuổi; Tiêu thức số người trong một hộ gia đình (1, 2, 3, 4,...) là tiêu thức với lượng biến rời rạc, tiêu thức doanh thu của doanh nghiệp, tiêu thức điểm số (0, 1, 2..., 10) là tiêu thức với lượng biến liên tục.

- *Tiêu thức thay phiên*: Là tiêu thức chỉ có 2 biểu hiện không trùng nhau trên một đơn vị tổng thể. Tiêu thức thay phiên có thể là tiêu thức thuộc tính và tiêu thức số lượng. Ví dụ: Là tiêu thức thuộc tính: tiêu thức giới tính chỉ có hai biểu hiện: nam và nữ; là tiêu thức số lượng: Tiêu thức năng suất lao động tiên tiến chỉ có hai biểu hiện: chẳng hạn NSLĐ >2000 sản phẩm/h là tiên tiến, < 200 sản phẩm/h là không tiên tiến;

Các tiêu thức thuộc tính, số lượng đều góp phần vào việc xác định đơn vị tổng thể cũng như tổng thể, nhờ đó chúng ta có thể phân biệt đơn vị này với đơn vị khác, tổng thể này với tổng thể khác.

1.2.3. Chỉ tiêu thống kê

1.2.3.1. Khái niệm

Chỉ tiêu thống kê phản ánh đặc điểm về mặt lượng của nhiều đơn vị, hiện tượng cá biệt của hiện tượng số lớn trong điều kiện thời gian và không gian cụ thể. Thông qua đó có thể thấy được nội dung, ý nghĩa, bản chất, tính quy luật của hiện tượng nghiên cứu. Đây là đặc điểm phân biệt giữa tiêu thức và chỉ tiêu thống kê. Tính chất của hiện tượng được khái quát hoá trong chỉ tiêu thống kê.

Như vậy: Chỉ tiêu thống kê là khái niệm biểu hiện mặt lượng trong mối liên hệ mật thiết với mặt chất của hiện tượng số lớn trong điều kiện thời gian và không gian cụ thể.

Trong Khoản 3, Điều 3 Luật Thống kê (được Quốc hội thông qua ngày 17/6/2003, có hiệu lực thi hành từ ngày 01/01/2004) đã quy định cụ thể hơn: *“Chỉ tiêu thống kê là tiêu chí mà biểu hiện bằng số của nó phản ánh quy mô, tốc độ phát triển, cơ cấu, quan hệ tỷ lệ của hiện tượng kinh tế - xã hội trong điều kiện không gian và thời gian cụ thể”.*

Kết cấu của một chỉ tiêu thống kê gồm 2 phần: Phần nội dung là tên gọi của chỉ tiêu quy định về khái niệm, giới hạn về thuộc tính, số lượng, thời gian của hiện tượng. Phần mức độ (hay con số) của chỉ tiêu phản ánh quy mô, tỷ lệ, quan hệ so sánh với đơn vị tính phù hợp.

Ví dụ: Tổng số dân Việt Nam vào 0h ngày 1/4/2009 là 85.846.997 người. Doanh thu của công ty X quý I năm 2015 là 80 tỷ đồng. Năng suất lao động bình quân đầu người của nhà máy A năm 2015 là 8000 m²/ người.

1.2.3.2. Phân loại chỉ tiêu thống kê

+ Căn cứ vào nội dung phản ánh, chia chỉ tiêu thành hai loại:

Chỉ tiêu chất lượng: Biểu hiện các tính chất, trình độ phổ biến, mối quan hệ của tổng thể. Chỉ tiêu chất lượng có thể biểu hiện bằng chỉ tiêu tương đối, chỉ tiêu trung bình. Ví dụ: Năng suất lao động, chi phí sản xuất/1 đơn vị sản phẩm, tỷ suất chi phí lưu thông, số bác sỹ/1000 dân.

Chỉ tiêu khối lượng: Biểu hiện quy mô của tổng thể. Ví dụ: Số nhân khẩu, số công nhân, số lượng sản phẩm sản xuất ra, doanh thu của doanh nghiệp, tổng thu nhập quốc dân.

+ *Căn cứ vào biểu hiện của các chỉ tiêu có thể phân biệt:*

Chỉ tiêu hiện vật: Biểu hiện bằng đơn vị hiện vật, có thể là đơn vị tự nhiên, đơn vị đo lường hoặc đơn vị đo lường quy ước.

Chỉ tiêu giá trị: Biểu hiện bằng đơn vị tiền tệ.

Chỉ tiêu hiện vật biểu hiện cụ thể hình thái vật chất của đối tượng nghiên cứu, nhưng khả năng tổng hợp hạn chế vì đơn vị tính khác nhau, còn chỉ tiêu giá trị có khả năng tính toán tổng hợp tuy nhiên thường bị ảnh hưởng bởi sự biến động của yếu tố giá. Đây là vấn đề cần lưu ý khi sử dụng các chỉ tiêu này trong phân tích thống kê.

1.2.4. Hệ thống chỉ tiêu thống kê

Hệ thống chỉ tiêu thống kê là tập hợp nhiều chỉ tiêu thống kê có quan hệ với nhau, được sắp xếp theo một thứ tự nào đó, phản ánh nhiều mặt của hiện tượng hay quá trình kinh tế - xã hội trong điều kiện không gian và thời gian cụ thể. Có nhiều loại hệ thống chỉ tiêu thống kê: Hệ thống chỉ tiêu thống kê quốc gia; Hệ thống chỉ tiêu thống kê của ngành, lĩnh vực; Hệ thống chỉ tiêu thống kê doanh nghiệp...

1.2.5. Dữ liệu thống kê

1.2.5.1. Khái niệm dữ liệu thống kê

Dữ liệu thống kê là các số liệu thu thập, tổng hợp và phân tích để phát hiện nội dung và ý nghĩa của nó. Tất cả dữ liệu có được trong một nghiên cứu cụ thể gọi là một bộ dữ liệu. Dữ liệu có thể là số đơn vị tổng thể hoặc là trị số của tiêu thức, chỉ tiêu.

1.2.5.2. Phân loại dữ liệu thống kê

- *Căn cứ vào tính chất*, dữ liệu thống kê được chia ra hai loại: Dữ liệu định tính và dữ liệu định lượng.

+ Dữ liệu định tính bao gồm tên hay nhãn được sử dụng để xác định đặc trưng của mỗi đơn vị hoặc phần tử. Dữ liệu định tính được đo

lượng bằng thang định danh hoặc thang thứ bậc và có thể được ký hiệu bằng con số hoặc các ký tự.

+ Dữ liệu định lượng là dữ liệu được biểu hiện bằng những con số cụ thể. Đối với dữ liệu định lượng có thể áp dụng các phép tính số học như cộng, trừ, nhân, chia để có được những thông tin có ý nghĩa hơn.

- *Căn cứ vào nguồn hình thành*, dữ liệu thống kê được chia thành hai loại: Dữ liệu thứ cấp và dữ liệu sơ cấp.

+ Dữ liệu thứ cấp là dữ liệu có sẵn trong một nghiên cứu thống kê cụ thể. Ví dụ: Khi phân tích kết quả sản xuất kinh doanh của một doanh nghiệp, có thể sử dụng các số liệu về doanh thu, lợi nhuận trong các báo cáo tài chính của bộ phận kế toán.

+ Dữ liệu sơ cấp là dữ liệu không có sẵn mà phải thông qua việc điều tra khảo sát mới có được.

1.3. Thang đo trong thống kê

Tuỳ theo tính chất của dữ liệu thống kê thu được, ta có thể sử dụng các loại thang đo khác nhau. Có bốn loại thang đo chủ yếu:

1.3.1. Thang đo định danh

Thang đo định danh là loại thang đo sử dụng cho các tiêu thức thuộc tính, mà các biểu hiện của dữ liệu không có sự hơn kém, khác biệt về thứ bậc, không theo một trật tự xác định nào như: Giới tính, khu vực địa lý, nghề nghiệp, tôn giáo... Trong trường hợp này các biểu hiện cùng loại của tiêu thức được đánh số như nhau.

Ví dụ: Tiêu thức giới tính có hai biểu hiện, có thể đánh số Nam: 1, Nữ: 2 và ngược lại.

Giữa các con số này không có quan hệ hơn kém, không thực hiện được các phép tính thống kê, mà chỉ đếm được tần số xuất hiện từng biểu hiện. Các dữ liệu định tính thường được đo lường bằng loại thang đo này.

1.3.2. Thang đo thứ bậc

Thang đo thứ bậc là thang đo định danh, nhưng giữa các biểu hiện của nó có thứ bậc hơn kém. Thang đo này được sử dụng cho tiêu thức

thuộc tính, mà các biểu hiện của dữ liệu có sự hơn kém, khác nhau về thứ bậc.

Ví dụ: Trình độ nghề nghiệp của người công nhân được chia theo bậc: 1, 2, 3, 4...; chất lượng các loại sản phẩm được phân loại: I, II, III... Những con số này thể hiện trật tự hơn kém giữa các biểu hiện của tiêu thức.

Thang đo này có ưu điểm là cho thấy sự khác biệt, sự hơn kém giữa các biểu hiện của tiêu thức, nhưng sự hơn kém này mới chỉ là tương đối vì chênh lệch giữa các biểu hiện của tiêu thức không nhất thiết phải bằng nhau. Cũng chính vì vậy không thực hiện được các phép tính số học trên thang đo này, mà chỉ thể hiện đặc trưng hơn hay kém của tổng thể một cách tương đối.

1.3.3. Thang đo khoảng

Thang đo khoảng là thang đo thứ bậc có khoảng cách bằng nhau, nhưng không có điểm gốc là 0. Ví dụ như: Thang đo nhiệt độ không khí...

Nhờ có tiêu chuẩn đo được quy định chính xác (với khoảng cách đều nhau), ta có thể đánh giá được mức độ hơn kém cụ thể về mặt lượng. Thang đo này luôn có đơn vị đo và được sử dụng cho tiêu thức số lượng. Từ đó, ta có thể thực hiện được các phép tính cộng, trừ, nhân, chia và tính được các tham số thống kê như trung bình, phương sai...

Hạn chế là thang đo này không có gốc 0 trên thực tế, mà chỉ có những điểm xác định các khoảng theo trật tự nào đó, nếu có điểm 0 chỉ là quy ước.

1.3.4. Thang đo tỷ lệ

Thang đo tỷ lệ là thang đo khoảng với giá trị 0 tuyệt đối (một trị số có thực), được coi như là điểm xuất phát của độ dài đo lường trên thang. Do có gốc 0, nên có thể so sánh được tỷ lệ giữa các số đo.

Đây là thang đo chặt chẽ nhất, được sử dụng phổ biến trong nghiên cứu thống kê. Với thang đo này có thể thực hiện được tất cả các công cụ toán, thống kê để tính toán và phân tích dữ liệu.

Trong bốn loại thang đo nói ở trên, thang đo sau có chất lượng đo lường cao hơn thang đo trước. Hai loại thang đo đầu được dùng để đo lường các dấu hiệu định tính và thường gọi là thang đo định tính. Thang đo khoảng và thang tỷ lệ là thang đo định lượng. Tuy nhiên, không phải sử dụng thang đo chất lượng cao hơn là tốt hơn, mà phải tùy thuộc vào đặc điểm của hiện tượng và tiêu thức nghiên cứu để sử dụng thang đo phù hợp.

1.4. Quá trình nghiên cứu thống kê

Một nghiên cứu thống kê thường phải trải qua một quá trình gồm nhiều giai đoạn, nhiều bước công việc kế tiếp nhau quan hệ chặt chẽ với nhau. Có thể chia quá trình này thành ba giai đoạn: Điều tra thống kê, tổng hợp thống kê, phân tích và dự báo thống kê.

- *Điều tra thống kê*: Điều tra thống kê là việc tổ chức một cách khoa học, theo một kế hoạch thống nhất việc thu thập, ghi chép nguồn tài liệu ban đầu về hiện tượng nghiên cứu ở một điều kiện cụ thể về thời gian và không gian. Đây là giai đoạn đầu tiên của quá trình nghiên cứu thống kê nhằm thu thập thông tin ban đầu cần thiết cho việc nghiên cứu. Để thực hiện việc thu thập thông tin, phải vận dụng nhiều hình thức, nhiều phương pháp khác nhau, tùy thuộc vào mục đích nghiên cứu, đặc điểm của hiện tượng nghiên cứu và điều kiện cụ thể của người nghiên cứu.

- *Tổng hợp thống kê*: Tổng hợp thống kê là quá trình tập trung, sắp xếp, hệ thống hóa các tài liệu ban đầu đã thu thập được ở giai đoạn điều tra thống kê. Nhiệm vụ của giai đoạn này là làm cho các đặc trưng riêng biệt của từng đơn vị điều tra bước đầu chuyển thành đặc trưng chung của toàn bộ tổng thể nghiên cứu, làm cơ sở cho giai đoạn phân tích và dự báo tiếp theo. Do đối tượng nghiên cứu của thống kê là hiện tượng số lớn, được cấu thành từ nhiều bộ phận có đặc điểm khác nhau nên việc tổng hợp được thực hiện theo từng tổ, từng nhóm. Vì vậy, phương pháp phân tổ là phương pháp cơ bản của giai đoạn này. Tài liệu sau khi tổng hợp được trình bày bằng bảng thống kê hoặc đồ thị thống kê.

- *Phân tích và dự báo thống kê*: Phân tích và dự báo thống kê là việc vận dụng các phương pháp phân tích để nêu lên một cách tổng hợp

nội dung, ý nghĩa, bản chất và tính quy luật qua các biểu hiện về mặt lượng của hiện tượng nghiên cứu. Trên cơ sở phân tích có thể xác định mức độ của hiện tượng nghiên cứu trong tương lai và đưa ra các căn cứ cho quyết định quản lý. Phân tích và dự báo là giai đoạn cuối cùng của quá trình nghiên cứu thống kê, biểu hiện tập trung kết quả của toàn bộ quá trình nghiên cứu thống kê. Phân tích thống kê giúp ta nhận thức được bản chất, tính quy luật của hiện tượng trong quá khứ, hiện tại và tương lai. Đồng thời giúp chỉ rõ mối liên hệ nội tại giữa các bộ phận của tổng thể, cũng như mối liên hệ, tác động qua lại lẫn nhau giữa hiện tượng nghiên cứu với các hiện tượng liên quan.

1.5. Tổ chức thống kê ở Việt Nam

1.5.1. Hệ thống tổ chức thống kê nhà nước ở Việt Nam

Theo quy định của Luật Thống kê và các văn bản hướng dẫn thi hành luật, hệ thống thống kê ở Việt Nam được tổ chức theo 2 hình thức: Thống kê nhà nước, Thống kê của các tổ chức, cá nhân ngoài hệ thống tổ chức thống kê nhà nước.

- Hệ thống tổ chức thống kê nhà nước:

Hệ thống tổ chức thống kê tập trung được tổ chức theo ngành dọc, gồm cơ quan thống kê trung ương và các cơ quan thống kê địa phương. Ở cấp trung ương có cơ quan Tổng cục thống kê. Tại các tỉnh, thành phố có các Cục thống kê tỉnh, thành phố trực thuộc Tổng cục thống kê. Cấp quận, huyện có các Phòng thống kê trực thuộc Cục thống kê. Những cơ quan này chịu trách nhiệm thực hiện nghĩa vụ, quyền hạn của cơ quan quản lý nhà nước ở từng cấp về mặt thống kê. Tại các xã, phường có các cán bộ làm công tác thống kê chịu sự quản lý trực tiếp của Ủy ban nhân dân cùng cấp và sự hướng dẫn về chuyên môn, nghiệp vụ của Phòng thống kê.

Các Bộ, cơ quan ngang Bộ, cơ quan thuộc Chính phủ, Tòa án nhân dân tối cao, Viện kiểm sát nhân dân tối cao đều phải có bộ phận thống kê (thường là phòng thống kê) thực hiện công tác thống kê theo quy định của Luật Thống kê. Cơ quan này do Thủ trưởng các Bộ, cơ quan ngang Bộ... tổ chức và quyết định.

- *Hệ thống thống kê các đơn vị cơ sở:*

Các doanh nghiệp, cơ quan hành chính, đơn vị sự nghiệp, đơn vị thuộc lực lượng vũ trang có tổ chức thống kê hoặc bố trí cán bộ thống kê đặt dưới sự quản lý của thủ trưởng đơn vị. Thống kê của các cơ quan này có trách nhiệm tổ chức thực hiện công tác thống kê phục vụ yêu cầu quản lý của đơn vị mình, phải gửi báo cáo tài chính cho Cục thống kê tỉnh, thực hiện các cuộc điều tra thống kê và chế độ báo cáo thống kê nhà nước theo quy định.

1.5.2. Chức năng, nhiệm vụ của cơ quan thống kê nhà nước Việt Nam

Nghị định 101/2003/NĐ-CP ngày 03/09/2003 của Chính phủ quy định chức năng, nhiệm vụ, quyền hạn và cơ cấu tổ chức của Tổng cục thống kê đã quy định: Tổng cục thống kê là cơ quan thuộc Chính phủ, có chức năng thực hiện một số nhiệm vụ, quyền hạn quản lý Nhà nước về thống kê, tổ chức thực hiện hoạt động thống kê và cung cấp thông tin thống kê kinh tế - xã hội cho các cơ quan, tổ chức và cá nhân theo quy định của pháp luật. Như vậy, cơ quan thống kê nhà nước Việt Nam có hai chức năng:

- Quản lý Nhà nước về thống kê trong toàn bộ các hoạt động kinh tế - xã hội của đất nước.

- Tổ chức thực hiện hoạt động thống kê, sản xuất và cung cấp thông tin thống kê kinh tế - xã hội cho các cơ quan, tổ chức và cá nhân theo quy định của pháp luật.

Nghị định cũng quy định sáu nhiệm vụ chủ yếu có liên quan đến nghiệp vụ thống kê là:

- Soạn thảo các văn bản quy phạm pháp luật về thống kê, các chiến lược, quy hoạch, kế hoạch về thống kê...

- Trình Thủ tướng Chính phủ ban hành hệ thống chỉ tiêu thống kê quốc gia, chế độ báo cáo thống kê cơ sở, thống kê tổng hợp, chương trình điều tra thống kê...

- Chỉ đạo, hướng dẫn và thẩm định về chuyên môn thống kê đối với các Bộ, cơ quan ngang Bộ, các tỉnh, thành phố...

- Tổ chức thu thập thông tin thống kê kinh tế - xã hội, xây dựng và quản lý cơ sở dữ liệu quốc gia về thông tin thống kê kinh tế - xã hội.

- Báo cáo Chính phủ các số liệu thống kê tổng hợp tình hình thực hiện các chương trình, mục tiêu quốc gia, các kế hoạch nhà nước và các báo cáo phân tích, dự báo về tình hình phát triển kinh tế - xã hội hàng năm.

- Công bố và cung cấp thông tin thống kê kinh tế - xã hội cho người sử dụng.

TÓM TẮT CHƯƠNG 1

Thống kê học là môn khoa học ra đời khoảng 300 năm trước nhưng có nguồn gốc lịch sử phát triển lâu dài. Thống kê ra đời và phát triển từ nhu cầu thực tiễn của xã hội, đó là một quá trình đúc rút kinh nghiệm trong lĩnh vực tính toán số liệu từ đơn giản đến phức tạp, hoàn thiện dần theo lịch sử phát triển của xã hội và trở thành môn khoa học độc lập như ngày nay. Hiện nay thống kê được coi là một công cụ quản lý vĩ mô quan trọng, đồng thời là công cụ quản lý hoạt động sản xuất kinh doanh của đơn vị cơ sở.

Đối tượng nghiên cứu của thống kê là *“mặt lượng trong sự liên hệ mật thiết với mặt chất của các hiện tượng kinh tế - xã hội số lớn, trong điều kiện thời gian và địa điểm cụ thể”*. Khoa học thống kê sử dụng một số khái niệm cơ bản bao gồm: Tổng thể thống kê, tiêu thức, chỉ tiêu, hệ thống chỉ tiêu. Để phục vụ cho việc đo lường dữ liệu định tính và định lượng, thống kê sử dụng bốn loại thang đo: Thang định danh, thang thứ bậc, thang đo khoảng, thang tỷ lệ.

Theo quy định trong Luật Thống kê của nước CHXHCN Việt Nam, hệ thống thống kê được tổ chức theo hai hình thức: Hệ thống thống kê tập trung, được tổ chức theo ngành dọc gồm cơ quan thống kê Trung ương và các cơ quan thống kê địa phương. Các Bộ, các cơ quan ngang Bộ có bộ phận thực hiện công tác thống kê theo luật định, ngoài ra các đơn vị cơ sở có bộ phận thống kê phục vụ yêu cầu quản lý của đơn vị và thực hiện chế độ báo cáo thống kê nhà nước theo quy định.

CÂU HỎI ÔN TẬP CHƯƠNG 1

1. Phân tích đối tượng nghiên cứu của thống kê học? Cho ví dụ minh họa?
2. Trình bày các khái niệm cơ bản của thống kê học, ý nghĩa của các khái niệm này? Cho ví dụ minh họa?
3. Phân biệt sự khác nhau của 2 khái niệm: Tiêu thức và Chỉ tiêu thống kê?
4. Trình bày các loại thang đo trong thống kê, cho ví dụ cụ thể về các trường hợp vận dụng các loại thang đo trên?
5. Trình bày cơ cấu tổ chức của hệ thống thống kê ở Việt Nam?

Chương 2

ĐIỀU TRA THỐNG KÊ

Đối tượng nghiên cứu thống kê là hiện tượng số lớn, là tập hợp các đơn vị thuộc tổng thể nghiên cứu. Quá trình thu thập tài liệu ban đầu thường được bắt đầu trên đơn vị gọi là điều tra thống kê, đây là giai đoạn đầu tiên của quá trình nghiên cứu thống kê. Do tính chất phức tạp của đối tượng cần thu thập dữ liệu, điều tra thống kê phải được tiến hành theo một phương pháp và quy trình tổ chức khoa học. Chương này sẽ trình bày những vấn đề lý luận, phương pháp, cách thức tổ chức để thu thập tài liệu ban đầu phục vụ cho một quá trình nghiên cứu thống kê.

2.1. Khái niệm, ý nghĩa và yêu cầu của điều tra thống kê

2.1.1. Khái niệm điều tra thống kê

Vấn đề đầu tiên của quá trình nghiên cứu thống kê là phải thu thập thông tin của từng đơn vị tổng thể. Điều tra thống kê chính là tiến hành thu thập tài liệu của từng đơn vị tổng thể trên cơ sở nội dung nghiên cứu.

Ví dụ: Để đánh giá về quy mô kinh doanh của các doanh nghiệp thì phải tiến hành thu thập tài liệu tại từng doanh nghiệp như: Vốn, tài sản, lao động, giá trị sản lượng... Trong Tổng điều tra dân số, các điều tra viên đến từng hộ gia đình để ghi chép các đặc điểm trên từng nhân khẩu. Hoặc như việc thăm dò ý kiến khách hàng, thông qua hình thức gửi các phiếu điều tra tới tay từng khách hàng để họ tự điền các thông tin vào phiếu. Các công việc tiến hành như trên gọi là điều tra thống kê.

Tuy nhiên, do số đơn vị cần thu thập tài liệu thường rất lớn, đa dạng và thường xuyên biến động, do vậy khi tiến hành điều tra thống kê đòi hỏi việc thu thập dữ liệu phải được tổ chức một cách khoa học, có kế hoạch thống nhất và chuẩn bị chu đáo. Vì vậy, có thể hiểu: *Điều tra thống kê là việc tổ chức một cách khoa học theo một kế hoạch thống nhất việc thu thập, ghi chép nguồn tài liệu ban đầu về hiện tượng nghiên cứu trong điều kiện cụ thể về thời gian, không gian.*

Điều 3, Luật Thống kê nước Cộng hoà xã hội chủ nghĩa Việt Nam cũng định nghĩa: “*Điều tra thống kê là hình thức thu thập thông tin thống kê theo phương án điều tra*”. Định nghĩa này hoàn toàn phù hợp với khái niệm nêu trên, bởi lẽ phương án điều tra thống kê sẽ quy định rõ về mục đích, ý nghĩa, toàn bộ quá trình tổ chức, điều kiện thời gian và không gian của cuộc điều tra.

2.1.2. Ý nghĩa của điều tra thống kê

Điều tra thống kê nếu được tổ chức một cách khoa học, sẽ giúp giải quyết được nhiều vấn đề trong nghiên cứu cũng như trong hoạt động thực tiễn. Cụ thể:

- Là căn cứ tin cậy để kiểm tra, đánh giá thực trạng hiện tượng nghiên cứu, đánh giá tình hình thực hiện kế hoạch phát triển kinh tế, văn hoá, xã hội của từng đơn vị, từng địa phương và của toàn bộ nền kinh tế quốc dân.

- Cung cấp những luận cứ xác đáng cho việc phân tích, phát hiện những yếu tố tác động, những yếu tố quyết định sự biến đổi của hiện tượng nghiên cứu, trên cơ sở đó tìm biện pháp thích hợp thúc đẩy hiện tượng phát triển theo hướng có lợi nhất.

- Căn cứ cho việc phát hiện, xác định xu hướng, quy luật biến động của hiện tượng và dự đoán xu hướng biến động của hiện tượng trong tương lai.

Theo cách thức tổ chức các hoạt động thống kê nhà nước Việt Nam, điều tra thống kê được tổ chức thành 2 cấp độ: Tổng điều tra thống kê và điều tra thống kê.

Tổng điều tra thống kê để thu thập những thông tin thống kê cơ bản trên phạm vi cả nước theo chu kỳ dài, quy mô lớn, phạm vi rộng liên quan đến nhiều ngành, nhiều cấp, có sử dụng nguồn kinh phí rất lớn.

Điều tra thống kê được thực hiện đối với các tổ chức, hộ gia đình và cá nhân trong các trường hợp sau:

- Thu thập những thông tin thống kê từ các tổ chức không phải thực hiện chế độ báo cáo thống kê định kỳ.

- Khi cần bổ sung thông tin từ các tổ chức có thực hiện chế độ báo cáo thống kê.

- Thu thập thông tin thống kê khi có nhu cầu đột xuất.

2.1.3. Những yêu cầu cơ bản của điều tra thống kê

Điều tra thống kê cần đảm bảo ba yêu cầu sau: Chính xác, kịp thời và đầy đủ.

- *Chính xác* trong điều tra thống kê có nghĩa là tài liệu thu thập được phải phản ánh trung thực tình hình thực tế khách quan, có sao ghi chép vậy không được tự tiện thêm bớt. Muốn vậy, nhân viên điều tra phải ghi chép trung thực, có tinh thần trách nhiệm và trình độ chuyên môn, tài liệu thu thập được phải theo đúng nội dung và đảm bảo đầy đủ về số lượng chỉ tiêu, số đơn vị tổng thể. Một tài liệu điều tra chính xác mới có thể dùng làm căn cứ tin cậy cho giai đoạn tổng hợp và phân tích tiếp theo.

- *Kịp thời* trong điều tra thống kê có nghĩa là thống kê phải nhạy bén với tình hình, thu thập và phản ánh đúng lúc các tài liệu mà lãnh đạo quan tâm, tiến hành đúng thời hạn quy định để phát huy hết tác dụng của tài liệu điều tra.

- *Đầy đủ* có nghĩa là được thu thập theo đúng nội dung điều tra đã quy định, đầy đủ số đơn vị điều tra đã được quy định trong phương án điều tra.

2.2. Các loại điều tra thống kê

2.2.1. Điều tra thường xuyên và không thường xuyên

Căn cứ vào tính chất liên tục, tính hệ thống của các cuộc điều tra, có thể chia làm 2 loại: Điều tra thường xuyên và điều tra không thường xuyên.

- *Điều tra thường xuyên*: Là tiến hành ghi chép, thu thập tài liệu ban đầu của hiện tượng một cách liên tục, có hệ thống và thường là theo sát với quá trình phát sinh, phát triển của hiện tượng.

Ví dụ: Ghi chép số sản phẩm sản xuất, doanh thu bán hàng, chi phí sản xuất, khối lượng hàng hoá tiêu thụ... Hoặc theo dõi nhân viên có mặt nơi làm việc thông qua bảng chấm công hàng ngày. Hay như đối với mỗi cửa hàng, đầu ngày và cuối ngày phải ghi chép lượng hàng hoá tồn kho, nhập kho, xuất kho trong ngày.

Điều tra thường xuyên có khả năng theo dõi tỉ mỉ tình hình phát triển của hiện tượng theo thời gian, thu thập được những tài liệu phản ánh quá trình phát triển tích lũy của hiện tượng trong thời kỳ dài. Tài liệu điều tra thường xuyên là cơ sở chủ yếu để lập báo cáo thống kê định kỳ, là công cụ quan trọng để theo dõi tình hình thực hiện kế hoạch.

- *Điều tra không thường xuyên*: Là thu thập tài liệu của hiện tượng nghiên cứu một cách không liên tục, mà chỉ tiến hành ghi chép tài liệu vào một thời điểm nào đó, không gắn liền với quá trình phát sinh phát triển của hiện tượng. Tài liệu của điều tra không thường xuyên phản ánh trạng thái của hiện tượng ở một thời điểm nhất định.

Loại điều tra này thường dùng cho các hiện tượng cần theo dõi thường xuyên nhưng chi phí điều tra lớn (điều tra dân số, điều tra nông nghiệp, tài sản cố định...) hoặc không xảy ra thường xuyên (điều tra dư luận...).

Ví dụ: Tổng điều tra dân số không thể điều tra hàng năm được mà chỉ định kỳ 10 năm một lần. Hay như không thể ngày nào cũng ghi chép được lượng hàng tồn kho mà thường chỉ kiểm kê, theo dõi vào cuối tuần hoặc cuối tháng hoặc cuối quý.

Cả hai loại điều tra thường xuyên và không thường xuyên đều có ưu điểm và nhược điểm riêng, do vậy tùy theo điều kiện thực tế mà nhất là kinh phí của mỗi một ngành, mỗi một quốc gia mà lựa chọn loại điều tra thích hợp và cũng có thể kết hợp cả hai loại điều tra cho cùng một hiện tượng.

Ví dụ: Ở nước ta, tiến hành Tổng điều tra dân số 10 năm một lần, nhưng vẫn tổ chức theo dõi biến động dân số thường xuyên thông qua khai sinh, chứng tử. Hoặc tổ chức những cuộc điều tra không định kỳ theo những mục đích khác nhau, như điều tra mức sinh, điều tra di dân.

2.2.2. Điều tra toàn bộ và không toàn bộ

Căn cứ vào phạm vi của đối tượng được điều tra thực tế, điều tra thống kê được chia thành: Điều tra toàn bộ và điều tra không toàn bộ.

- *Điều tra toàn bộ*: Là tiến hành thu thập tài liệu ban đầu trên toàn bộ các đơn vị thuộc đối tượng điều tra, không loại trừ bất kỳ đơn vị nào.

Ví dụ: Trong các cuộc tổng điều tra dân số, việc thu thập thông tin được tiến hành trên tất cả các công dân Việt Nam cư trú thường xuyên trên lãnh thổ Việt Nam có tại thời điểm điều tra, như vậy đây là cuộc điều tra toàn bộ.

Điều tra toàn bộ là nguồn cung cấp tài liệu đầy đủ nhất cho nghiên cứu thống kê. Nó thu thập tài liệu của toàn bộ các đơn vị thuộc đối tượng nghiên cứu, do đó ta có thể tính được các chỉ tiêu tổng hợp cho toàn bộ tổng thể, cũng như từng bộ phận trong đó. Tuy nhiên, điều tra toàn bộ đòi hỏi nguồn kinh phí rất lớn, vì vậy không thể làm thường xuyên được. Đối với những tổng thể tiềm ẩn, do không thể nhận biết được tất cả các đơn vị của tổng thể nghiên cứu, nên dễ bị bỏ sót trong quá trình điều tra nếu tiến hành điều tra toàn bộ.

- *Điều tra không toàn bộ*: Là tiến hành thu thập tài liệu ban đầu trên một số đơn vị được chọn ra trong toàn bộ các đơn vị của tổng thể chung. Những đơn vị được chọn phải đảm bảo một số điều kiện nhất định.

Ví dụ: Khi điều tra mức sống dân cư một địa phương thường chỉ chọn một số gia đình để tiến hành điều tra thực tế.

Điều tra không toàn bộ có ưu điểm là có thể giảm chi phí điều tra, tiến hành nhanh chóng, do vậy đảm bảo yêu cầu kịp thời, và có thể đi sâu vào phân tích nhiều khía cạnh của hiện tượng nghiên cứu. Do vậy, đây là loại điều tra được sử dụng rất phổ biến trong nghiên cứu thống kê.

Tuỳ theo cách lựa chọn các đơn vị để điều tra, có thể phân biệt các loại điều tra không toàn bộ sau đây:

+ *Điều tra chọn mẫu*: Là một loại điều tra không toàn bộ, trong đó chỉ chọn ra một số đơn vị để điều tra thực tế, sau đó dùng kết quả

thu thập được để tính toán suy rộng thành các đặc điểm của toàn bộ tổng thể.

Các đơn vị mẫu được chọn theo một nguyên tắc nhất định để đảm bảo tính chất đại biểu đại diện cho tổng thể nghiên cứu.

Trong thực tế đây là loại điều tra được sử dụng rất phổ biến vì những ưu điểm của nó. Loại điều tra này có thể thay thế điều tra toàn bộ khi chưa có điều kiện điều tra toàn bộ, bởi vì chỉ điều tra trên một số ít đơn vị nhưng kết quả lại có thể suy rộng cho tất cả các đơn vị tổng thể.

Ví dụ:

+ Trong sản xuất công nghiệp, có thể dùng điều tra chọn mẫu để kiểm nghiệm chất lượng sản phẩm, nghiên cứu tình hình năng suất lao động và tiền lương công nhân, tình hình sử dụng lao động, tình hình sử dụng máy móc thiết bị và nguyên vật liệu...

+ Để biết tình hình kinh doanh trong ngành thương mại không thể điều tra ở tất cả các doanh nghiệp thương mại, mà chỉ chọn ra một số doanh nghiệp để điều tra từ đó suy rộng kết quả cho toàn ngành.

+ Trong điều tra năng suất và sản lượng nông nghiệp, người ta lựa chọn một số đơn vị diện tích gieo cấy của địa phương, theo nguyên tắc mỗi đơn vị đều có khả năng được chọn như nhau, không kể lúa trên thửa ruộng được chọn là tốt hay xấu, sau khi điều tra sản lượng trên các đơn vị được chọn, người ta căn cứ vào kết quả đó để suy rộng ra năng suất và sản lượng lúa của toàn địa phương.

Nội dung cụ thể của loại điều tra này sẽ được đi sâu nghiên cứu trong chương 8.

+ **Điều tra trọng điểm:** Là loại điều tra không toàn bộ, trong đó chỉ tiến hành thu thập tài liệu trên những bộ phận chủ yếu, trọng điểm của hiện tượng nghiên cứu. Kết quả điều tra không suy rộng mà chỉ tính toán, phân tích, đánh giá những đặc điểm chung của tổng thể, giúp ta nhận thức được tình hình cơ bản của hiện tượng nghiên cứu.

Điều tra trọng điểm thích hợp với những đối tượng có từng bộ phận tương đối tập trung, bộ phận này chiếm tỷ trọng tương đối lớn trong tổng thể.

+ *Điều tra chuyên đề*: Là loại điều tra không toàn bộ, chỉ tiến hành thu thập tài liệu trên một số rất ít đơn vị nhưng lại đi sâu nghiên cứu chi tiết nhiều khía cạnh khác nhau của đơn vị đó nhằm rút ra vấn đề đặc thù, cốt lõi. Tài liệu thu thập được trong điều tra chuyên đề không dùng để suy rộng hoặc làm căn cứ đánh giá tình hình cơ bản của hiện tượng nghiên cứu mà chỉ tìm ra các đơn vị tiên tiến điển hình để nghiên cứu kinh nghiệm của các đơn vị đó hoặc nghiên cứu tìm hiểu những vấn đề mới phát sinh.

2.3. Phương pháp thu thập thông tin trong điều tra thống kê

Để thu thập thông tin trong điều tra thống kê, người ta có thể sử dụng nhiều phương pháp khác nhau. Tùy điều kiện thực tế và đặc điểm của hiện tượng nghiên cứu, khả năng về tài chính, thời gian, kinh nghiệm, trình độ của người tổ chức và điều tra viên, để lựa chọn phương pháp điều tra thích hợp. Sau đây là một số phương pháp chủ yếu trong điều tra thống kê:

2.3.1. Phương pháp đăng ký trực tiếp

Theo phương pháp này nhân viên điều tra phải trực tiếp tiếp xúc với đối tượng điều tra, trực tiếp tiến hành hoặc giám sát việc cân đong, đo, đếm và sau đó ghi chép những thông tin thu được vào phiếu điều tra.

Ví dụ: Khi điều tra tồn kho hàng hoá, vật tư, nhân viên điều tra trực tiếp cân, đong, đo, đếm số lượng vật tư, hàng hoá còn lại trong kho, sau đó ghi kết quả vào phiếu điều tra.

Phương pháp đăng ký trực tiếp thường được thực hiện gắn liền với quá trình phát sinh, phát triển của hiện tượng. Tài liệu ghi chép ban đầu được đăng ký trực tiếp có độ chính xác cao nhưng lại đòi hỏi nhiều nhân lực và thời gian. Mặt khác, trong thực tế nhiều hiện tượng không cho phép quan sát, đo, đếm trực tiếp quá trình phát sinh, phát triển của chúng. Vì vậy, phạm vi áp dụng của phương pháp này có nhiều hạn chế.

2.3.2. Phương pháp phỏng vấn

Đây là phương pháp được sử dụng phổ biến, theo đó việc ghi chép, thu thập tài liệu ban đầu được thực hiện thông qua quá trình hỏi - đáp giữa nhân viên điều tra và người cung cấp thông tin.

Trong điều tra thống kê, phỏng vấn không phải là cuộc nói chuyện thông thường, cũng không phải là một cuộc thẩm vấn của nhân viên điều tra. Phỏng vấn trong thống kê phải tuân thủ theo mục tiêu nghiên cứu, theo đối tượng, khách thể, nội dung nghiên cứu đã được xác định trước trong phương án điều tra. Do vậy, nhân viên điều tra phải có sự chuẩn bị trước về kỹ năng phỏng vấn, về năng lực chuyên môn.

Trong thực tế, phỏng vấn có thể thích ứng với nhiều hoàn cảnh khác nhau mà không cần theo sát quá trình phát sinh, phát triển của hiện tượng. Mặt khác, thông tin thu được qua phỏng vấn thường có độ tin cậy cao, dễ tổng hợp. Do đó, phỏng vấn được sử dụng rộng rãi nhất trong điều tra thống kê nhằm thu thập nguồn tài liệu ban đầu.

Nếu căn cứ vào tính chất của sự tiếp xúc giữa người hỏi và người trả lời, có thể chia ra: Phỏng vấn trực tiếp và phỏng vấn gián tiếp.

+ Phỏng vấn trực tiếp: Là phương pháp ghi chép, thu thập tài liệu ban đầu được thực hiện thông qua quá trình hỏi - đáp trực tiếp giữa nhân viên điều tra và người cung cấp thông tin. Tức là nhân viên điều tra trực tiếp đến địa bàn điều tra, tìm gặp đối tượng phỏng vấn, trực tiếp hỏi và ghi câu trả lời vào phiếu điều tra.

Do việc tiếp xúc trực tiếp giữa người hỏi và người trả lời nên phương pháp này tạo ra những điều kiện để hiểu đối tượng sâu sắc hơn. Giúp nhân viên điều tra hiểu thêm thông tin qua cử chỉ, thái độ của người trả lời, đồng thời có thể giải thích thêm nội dung của câu hỏi nếu người được hỏi chưa hiểu đầy đủ. Tuy nhiên, cách tiến hành này thường cần nhiều thời gian, chi phí, số nhân viên điều tra, đòi hỏi có sự chuẩn bị kỹ càng của nhân viên điều tra.

+ Phỏng vấn gián tiếp: Là phương pháp thu thập tài liệu ban đầu được thực hiện bằng cách: người được hỏi nhận được phiếu điều tra, tự mình ghi câu trả lời vào phiếu rồi gửi trả lại cho cơ quan điều tra.

Đặc điểm của phương pháp này là người hỏi và người trả lời không trực tiếp gặp nhau. Quá trình hỏi đáp diễn ra thông qua một trung gian là phiếu điều tra.

Cách thức tiến hành này có ưu điểm là dễ tổ chức, tiết kiệm chi phí và không cần nhiều điều tra viên. Nhược điểm cơ bản của nó là khó kiểm

tra, đánh giá được độ chuẩn xác của câu trả lời, tỷ lệ thu hồi phiếu trong nhiều trường hợp không cao, nội dung điều tra bị hạn chế. Phương pháp này chỉ có thể sử dụng trong điều kiện cộng đồng người được hỏi có trình độ dân trí cao.

Ngoài ra, trong điều tra thống kê, người ta còn sử dụng nhiều phương pháp khác để thu thập nguồn tài liệu ban đầu: Phương pháp quan sát, phương pháp thu thập thông tin qua nguồn sẵn có, phương pháp đăng ký qua chứng từ sổ sách...

2.4. Hình thức tổ chức điều tra thống kê

Theo đặc tính của hiện tượng nghiên cứu và yêu cầu nghiên cứu, có hai hình thức tổ chức điều tra:

2.4.1. Báo cáo thống kê định kỳ

Để có thể tổng hợp được dữ liệu của các chỉ tiêu thống kê thuộc một hệ thống chỉ tiêu thống kê nào đó và để tổng hợp các thông tin thống kê khác nhằm phục vụ cho việc quản lý nhà nước thì cơ quan nhà nước có thẩm quyền ban hành những quy định và hướng dẫn về việc thực hiện tập hợp các thông tin thống kê có liên quan, những văn bản quy định này được hiểu là những chế độ báo cáo thống kê, trong chế độ báo cáo thống kê luôn quy định và hướng dẫn cụ thể, thống nhất các nội dung có liên quan, các phương pháp thu thập thông tin cũng như biểu mẫu cho mỗi báo cáo thống kê định kỳ.

Báo cáo thống kê định kỳ là hình thức thu thập dữ liệu, thông tin thống kê về hiện tượng nghiên cứu trong một thời kỳ nhất định dựa trên các biểu mẫu báo cáo thống kê đã được lập sẵn do cơ quan nhà nước có thẩm quyền quy định. Đây là hình thức thu thập tài liệu mà theo đó các đơn vị báo cáo tiến hành ghi chép số liệu theo biểu mẫu có sẵn và gửi lên cấp trên theo quy định.

Các báo cáo thống kê được thực hiện một cách định kỳ thường xuyên, theo nội dung, phương pháp, biểu mẫu đã được quy định một cách khoa học và thống nhất. Trong hình thức này, người ta áp dụng phổ biến loại điều tra toàn bộ và thường xuyên. Đây là hình thức tổ chức điều tra theo con đường hành chính bắt buộc.

Báo cáo thống kê định kỳ bao gồm báo cáo thống kê quốc gia và báo cáo thống kê Bộ, Ngành:

Báo cáo thống kê quốc gia được thực hiện nhằm thu thập những thông tin thuộc hệ thống chỉ tiêu thống kê Quốc gia. Báo cáo thống kê quốc gia được thực hiện bởi các đơn vị, cơ quan như: Bộ, Ngành, cơ quan trung ương của tổ chức chính trị, xã hội, các cơ quan, tổ chức khác theo quy định của Pháp luật.

Báo cáo thống kê Bộ, Ngành được thực hiện nhằm thu thập những thông tin thống kê thuộc hệ thống chỉ tiêu thống kê quốc gia, hệ thống chỉ tiêu thống kê bộ, ngành, hệ thống chỉ tiêu thống kê cấp tỉnh, cấp huyện và tổng hợp các thông tin thống kê khác phục vụ yêu cầu quản lý ngành, lĩnh vực. Cơ quan, tổ chức thực hiện chế độ báo cáo thống kê cấp bộ, ngành gồm: Cơ quan chuyên môn thuộc Ủy ban nhân dân cấp tỉnh, cấp huyện; Tòa án nhân dân, Viện kiểm sát nhân dân cấp tỉnh, cấp huyện; cơ quan chuyên môn thuộc bộ, cơ quan ngang bộ đặt tại địa phương; Cơ quan, tổ chức khác theo quy định của pháp luật.

Nội dung cơ bản của báo cáo thống kê định kỳ bao gồm những chỉ tiêu cơ bản về hoạt động nghiệp vụ có liên quan chặt chẽ đến việc thực hiện nhiệm vụ kế hoạch đặt ra. Căn cứ vào nguồn tài liệu do các báo cáo phản ánh một cách có hệ thống, cơ quan lãnh đạo có thể thường xuyên và kịp thời chỉ đạo nghiệp vụ đối với cấp dưới, giám sát và kiểm tra tình hình thực hiện kế hoạch, phát hiện những tồn tại, yếu kém để có biện pháp khắc phục. Tài liệu của báo cáo thống kê định kỳ còn được dùng để phân tích, đối chiếu giữa các đơn vị rút ra những kết luận chung về vấn đề đang được nghiên cứu, theo dõi và là căn cứ để xây dựng kế hoạch cho kỳ sau.

2.4.2. Điều tra chuyên môn

Điều tra chuyên môn là hình thức tổ chức thu thập dữ liệu, thông tin thống kê về đối tượng nghiên cứu cụ thể theo phương pháp khoa học, thống nhất được quy định trong phương án điều tra thống kê cho mỗi lần điều tra. Như vậy, mỗi một cuộc điều tra sẽ có một phương án điều tra riêng, trong đó quy định rõ những vấn đề cần được hiểu thống nhất và hướng dẫn cụ thể, chi tiết việc thực hiện điều tra.

Điều tra chuyên môn phù hợp với các cuộc điều tra không thường xuyên, dùng cho những trường hợp chỉ khi có nhu cầu thông tin mới tổ chức tiến hành vào một thời điểm hoặc trong một thời kỳ nhất định hoặc dùng thu thập thông tin về những hiện tượng có sự thay đổi liên tục theo thời gian nhưng không nhất thiết phải theo dõi thường xuyên (như hiện tượng về dân số, tình hình sử dụng đất đai, hệ thống cơ sở hạ tầng...).

Nội dung của điều tra chuyên môn luôn thay đổi theo từng lần điều tra, nội dung trong điều tra chuyên môn thường là những hiện tượng mà báo cáo thống kê định kỳ chưa tiến hành hoặc không thể thu thập được. Hoặc đó là những vấn đề có liên quan đến kinh tế tư nhân, những hiện tượng xảy ra bất thường (như bão lụt, thiên tai). Trong một số trường hợp điều tra chuyên môn còn được tổ chức để thu thập lại những thông tin đã được ghi chép trong báo cáo thống kê nhằm kiểm tra chất lượng của báo cáo thống kê định kỳ, chỉnh sửa hoặc bổ sung những chi tiết mà báo cáo này chưa phản ánh được.

Trong điều kiện kinh tế thị trường ở nước ta hiện nay vai trò của điều tra chuyên môn ngày càng quan trọng vì nó đảm bảo tính hữu dụng, khắc phục được sự cứng nhắc của báo cáo thống kê định kỳ. Do đó, nó ngày càng được sử dụng rộng rãi để phục vụ cho nhiều yêu cầu nghiên cứu cũng như công tác thực tiễn khác nhau.

2.5. Xây dựng phương án điều tra

Để tiến hành một cuộc điều tra thống kê, cần thiết phải xây dựng một phương án điều tra.

Phương án điều tra thống kê là một loại văn bản trong đó quy định rõ những vấn đề cần phải giải quyết và những vấn đề cần được hiểu một cách thống nhất trước, trong và sau khi tiến hành điều tra. Để đảm bảo sự thống nhất trong việc tổ chức điều tra, tránh các sai sót trong quá trình điều tra thì phương án điều tra phải được xây dựng một cách chi tiết, tỷ mỉ, cụ thể và khoa học có sự phân công chức năng, nhiệm vụ rõ ràng giữa các đơn vị thực hiện. Văn bản này ngoài việc hướng dẫn nhân viên điều tra thực hiện cuộc điều tra, còn là cơ sở để các cấp có thẩm quyền quyết định và phê duyệt cho phép tiến hành cuộc điều tra.

Đối với cuộc điều tra lớn, có liên quan đến nhiều cấp, nhiều ngành thì phương án điều tra phải cần có sự phối hợp, bàn bạc thống nhất giữa các cơ quan trên.

Tùy thuộc vào quy mô, tính chất của từng cuộc điều tra mà phương án điều tra có thể có những sự khác nhau, nhưng nhìn chung một phương án điều tra gồm các nội dung chủ yếu: Mục đích, yêu cầu của điều tra; phạm vi, đối tượng, đơn vị điều tra; loại điều tra; thời điểm, thời gian và phương pháp điều tra; nội dung và phiếu điều tra; phân loại thống kê sử dụng trong điều tra; quy trình xử lý và biểu đầu ra của điều tra; tổ chức điều tra; kinh phí và các điều kiện vật chất cho điều tra.

2.5.1. Xác định mục đích điều tra

Bất kỳ một hiện tượng kinh tế - xã hội nào cũng có thể được quan sát, xem xét, nghiên cứu trên nhiều góc độ khác nhau, mỗi khía cạnh nghiên cứu sẽ giải quyết được những vấn đề cụ thể khác nhau. Vì vậy, trước khi tiến hành điều tra cần phải xác định rõ xem cuộc điều tra này nhằm tìm hiểu vấn đề gì, phục vụ cho yêu cầu nghiên cứu nào. Đó chính là mục đích của cuộc điều tra.

Mục đích điều tra còn là căn cứ quan trọng để xác định đối tượng, đơn vị điều tra, xây dựng kế hoạch và nội dung điều tra. Vì vậy, việc xác định đúng, rõ mục đích điều tra là cơ sở quan trọng cho việc thu thập số liệu ban đầu đầy đủ, hợp lý, đáp ứng yêu cầu nghiên cứu đặt ra.

Căn cứ để xác định mục đích điều tra thường là những nhu cầu thực tế trong đời sống xã hội hoặc phục vụ cho một yêu cầu nghiên cứu cụ thể nào đó.

Ví dụ: Mục đích của cuộc Tổng điều tra dân số năm 2009 là:

- Làm căn cứ chính xác phục vụ công tác nghiên cứu, phân tích quá trình phát triển kinh tế - xã hội và dân số nước ta.

- Kiểm điểm tình hình thực hiện các nhiệm vụ liên quan đến dân số, lao động. Trên cơ sở đó đáp ứng yêu cầu xây dựng kế hoạch phát triển kinh tế - xã hội thời kỳ 2011 - 2020.

2.5.2. Xác định đối tượng và đơn vị điều tra

Xác định đối tượng điều tra là xác định xem những đơn vị tổng thể nào thuộc phạm vi điều tra, cần được thu thập tài liệu. Xác định đối tượng điều tra là xác định được ranh giới giữa hiện tượng nghiên cứu với các tổng thể khác, tránh tình trạng trùng lặp hay bỏ sót khi tiến hành điều tra.

Muốn xác định chính xác đối tượng điều tra cần dựa vào phân tích lý luận, nêu lên những tiêu chuẩn cơ bản phân biệt hiện tượng nghiên cứu với hiện tượng có liên quan, đồng thời còn phải căn cứ vào mục đích nghiên cứu.

Đơn vị điều tra là đơn vị thuộc đối tượng điều tra và được điều tra thực tế. Đơn vị điều tra là nơi phát sinh các tài liệu ban đầu, điều tra viên đến đó để thu thập tài liệu.

Ví dụ: Trong điều tra tình hình phát triển các doanh nghiệp công nghiệp thì đối tượng điều tra là các doanh nghiệp trong ngành công nghiệp còn đơn vị điều tra đều là từng doanh nghiệp công nghiệp trên một địa bàn cụ thể. Còn trong Tổng điều tra dân số thì đối tượng điều tra được xác định là “tất cả công dân Việt Nam cư trú thường xuyên trên lãnh thổ Việt Nam vào thời điểm 0 giờ ngày 01 tháng 4 của năm tiến hành điều tra”, hay còn gọi là nhân khẩu thường trú. Đơn vị điều tra là các hộ gia đình.

Cần phân biệt đơn vị điều tra và đơn vị tổng thể. Đơn vị tổng thể là các phần tử cá biệt cấu thành hiện tượng cần được quan sát, còn đơn vị điều tra là đơn vị cung cấp thông tin, là nơi phát sinh tài liệu mà điều tra viên phải đến đó để thu thập tài liệu. Trong nhiều trường hợp đơn vị điều tra và đơn vị tổng thể không trùng nhau. Việc xác định đơn vị tổng thể liên quan đến việc lập phương án điều tra, ước lượng kinh phí điều tra, còn đơn vị điều tra chỉ liên quan đến việc tổ chức ghi chép tài liệu.

2.5.3. Chọn thời điểm, thời kỳ và quyết định thời hạn điều tra

+ *Thời điểm điều tra*: Là mốc thời gian được quy định thống nhất mà cuộc điều tra phải thu thập thông tin về hiện tượng tồn tại đúng thời điểm đó. Nếu việc ghi chép được tiến hành vào thời điểm sau đó, thì

người trả lời phải hồi tưởng lại để miêu tả trạng thái của hiện tượng vào đúng thời điểm điều tra.

+ *Thời kỳ điều tra*: Là khoảng thời gian (tuần, tháng năm...) được quy định để thu thập số liệu về hiện tượng được tích lũy trong cả thời kỳ đó.

+ *Thời gian điều tra hay thời hạn điều tra*: Là khoảng thời gian dành cho việc thực hiện nhiệm vụ thu thập số liệu. Thời hạn dài hay ngắn phụ thuộc vào quy mô, tính phức tạp của hiện tượng nghiên cứu và nội dung điều tra, khả năng, kinh nghiệm của điều tra viên. Nhìn chung thời hạn điều tra không nên quá cách xa thời điểm điều tra vì có thể làm mất thông tin do người trả lời không nhớ đầy đủ các sự kiện.

2.5.4. Xác định nội dung điều tra và thiết lập phiếu điều tra

Nội dung điều tra là toàn bộ các đặc điểm cơ bản của từng đối tượng, từng đơn vị điều tra mà ta cần thu được thông tin. Trong thực tế các đơn vị của hiện tượng nghiên cứu thường có rất nhiều đặc điểm khác nhau. Tuy nhiên, ta không thể và cũng không cần thiết phải thu thập thông tin của toàn bộ các tiêu thức đó, mà chỉ cần những tiêu thức có liên quan đến mục đích nghiên cứu, phục vụ được cho việc nghiên cứu. Vì vậy, bất kỳ cuộc điều tra nào cũng cần phải xác định rõ nội dung điều tra.

Việc xác định nội dung điều tra, cần căn cứ vào các yếu tố sau:

- Căn cứ vào mục đích điều tra vì mục đích điều tra đưa ra định hướng cần thu thập những thông tin nào để đáp ứng yêu cầu đó. Mục đích điều tra khác nhau, nhu cầu thông tin cũng khác nhau và do vậy nội dung điều tra cũng khác nhau, mục đích càng phong phú, nội dung điều tra càng bao gồm nhiều vấn đề.

- Đặc điểm của hiện tượng nghiên cứu: Hiện tượng nghiên cứu bao giờ cũng tồn tại những đặc điểm gắn liền điều kiện cụ thể về thời gian và không gian nhất định. Những đặc điểm đó liên quan đến việc xác định các nội dung cần thu thập trên đơn vị điều tra để đáp ứng được mục đích nghiên cứu. Khi thời gian và không gian thay đổi thì đặc điểm của hiện tượng có thể thay đổi theo, vì vậy tiêu thức cần thu thập cũng phải khác nhau.

- Năng lực, trình độ thực tế của đơn vị, của người tổ chức điều tra. Điều này biểu hiện ở khả năng về tài chính, về thời gian, kinh nghiệm, trình độ tổ chức điều tra.

Để có thể thu thập được thông tin chính xác và đầy đủ, nội dung của mỗi cuộc điều tra được diễn đạt cụ thể thành những câu hỏi ngắn gọn, rõ ràng, dễ hiểu được trình bày cụ thể trong phiếu điều tra. *Phiếu điều tra (hay còn gọi là biểu điều tra, bảng hỏi)* là tập hợp các câu hỏi phản ánh nội dung điều tra, được sắp xếp theo một trật tự logic nhất định.

- Tùy theo yêu cầu, nội dung và đối tượng, mỗi cuộc điều tra có thể phải xây dựng nhiều loại phiếu điều tra khác nhau.

Ví dụ: Trong Tổng điều tra dân số và nhà ở năm 2009, người ta xây dựng 2 loại phiếu: Phiếu điều tra toàn bộ và phiếu điều tra mẫu, phiếu điều tra toàn bộ sử dụng để thu thập thông tin trên tất cả các nhân khẩu là người Việt Nam, phiếu điều tra mẫu sử dụng trong điều tra cho một mẫu, phiếu điều tra mẫu ngoài việc thu thập những nội dung như trên phiếu điều tra toàn bộ còn bổ sung thêm một số nội dung về tình trạng hôn nhân, tình trạng khuyết tật, tình hình sinh, chết...Hoặc trong điều tra vốn đầu tư phát triển năm 2015, nội dung điều tra được thiết kế trên 8 phiếu điều tra áp dụng cho từng đơn vị điều tra riêng.

Thông thường trong phương án điều tra, cần có bản giải thích cách ghi phiếu điều tra nhằm giúp cho nhân viên điều tra và người trả lời nhận thức đúng về các câu hỏi, cách ghi chép số liệu. Đối với những vấn đề không cụ thể, khó trả lời, có nhiều cách hiểu khác nhau thì cần đưa ra ví dụ cụ thể và những quy định về các trường hợp ngoại lệ.

2.5.5. Các danh mục và bảng phân loại

Trong phương án điều tra luôn đưa ra các danh mục hoặc các bảng phân loại thống kê đã được xây dựng sẵn mà cuộc điều tra cần sử dụng. Các phân loại thống kê được sử dụng thống nhất trong hoạt động thống kê nhà nước, làm căn cứ thống nhất sử dụng trong quản lý nhà nước. Phân loại thống kê đóng vai trò quan trọng trong khâu xử lý và lập bảng số liệu, giúp chúng ta nhận thức được các hiện tượng kinh tế - xã hội một cách sâu sắc. Trong phương án điều tra các phân loại thống kê còn giúp

cho việc định hướng thu thập thông tin và phân công trách nhiệm cho các đơn vị thực hiện.

2.5.6. Loại điều tra và phương pháp thu thập thông tin

Tùy thuộc vào từng cuộc điều tra có thể sử dụng các loại điều tra khác nhau, có thể sử dụng riêng hoặc kết hợp các loại điều tra với nhau cho từng đối tượng điều tra hoặc đơn vị đã được xác định trước đó. Ví dụ, trong điều tra giá cước vận tải kho bãi thời kỳ 2010-2015 chỉ chọn điều tra chọn mẫu, còn trong điều tra vốn đầu tư phát triển năm 2015 được thực hiện theo 2 phương pháp là điều tra toàn bộ kết hợp với điều tra chọn mẫu, trong đó: Điều tra toàn bộ áp dụng với các đơn vị điều tra như UBND xã/phường/thị trấn cấp phép xây dựng các dự án/công trình mang tính xã hội hóa của các tổ chức/cá nhân đóng góp; sở tài chính tỉnh/thành phố; dự án được ủy quyền hoặc phân cấp thuộc Bộ/UBND; Hộ thuộc làng nghề sản xuất tập trung... điều tra không toàn bộ áp dụng với khu vực dân cư có đầu tư cho hoạt động sản xuất kinh doanh và nhà ở.

Trong phương án điều tra việc lựa chọn phương pháp thu thập thông tin phụ thuộc vào đặc điểm của hiện tượng nghiên cứu qua đó có thể chọn được phương pháp phù hợp, việc thu thập thông tin có thể được thực hiện bằng việc sử dụng một phương pháp hoặc có những trường hợp nội dung điều tra phức tạp bắt buộc phải sử dụng kết hợp nhiều phương pháp. Ví dụ, trong Tổng điều tra dân số và nhà ở năm 2009 việc thu thập thông tin được thực hiện bằng phương pháp điều tra trực tiếp, trong cuộc điều tra vốn đầu tư phát triển năm 2015 do tính chất phức tạp của nội dung điều tra và trình độ kế toán cũng như khả năng ứng dụng công nghệ thông tin của các đơn vị điều tra khác nhau, cuộc điều tra buộc phải sử dụng kết hợp hai phương pháp là điều tra trực tiếp và điều tra gián tiếp.

Trong mỗi cuộc điều tra, phương pháp thu thập thông tin trong phương án điều tra luôn phải được hướng dẫn một cách chi tiết, cụ thể cách thức tiến hành thu thập số liệu nhằm đảm bảo thông tin thu được chính xác, thống nhất.

2.5.7. Lập kế hoạch tổ chức và tiến hành điều tra

Lập kế hoạch tổ chức và tiến hành điều tra là một vấn đề trọng yếu của điều tra thống kê. Kế hoạch này quy định cụ thể từng bước công việc phải tiến hành trong quá trình từ khâu tổ chức đến khâu điều tra thực tế. Vì vậy, nó xây dựng càng chi tiết, tỷ mỉ, rõ ràng, cụ thể thì càng dễ thực hiện, chất lượng của cuộc điều tra càng nâng cao. Tuy nhiên, đây là một công việc phức tạp, đòi hỏi kinh nghiệm và am hiểu thực tế. Kế hoạch tổ chức bao gồm các khâu:

- Thành lập Ban chỉ đạo điều tra và quy định nhiệm vụ cụ thể cho cơ quan điều tra các cấp.

- Chuẩn bị lực lượng cán bộ điều tra, phân công trách nhiệm và tiến hành tập huấn nghiệp vụ cho họ.

- Lựa chọn phương pháp điều tra thích hợp.

- Định các bước tiến hành điều tra.

- Phân chia khu vực và địa bàn điều tra.

- Tổ chức các cuộc hội nghị chuẩn bị.

- Tiến hành điều tra thử nghiệm để rút kinh nghiệm, nâng cao trình độ nghiệp vụ cho cán bộ điều tra và hoàn thiện phương án điều tra, phiếu điều tra.

- Xây dựng phương án tài chính và chuẩn bị các phương tiện vật chất khác.

- Tuyên truyền mục đích, ý nghĩa của cuộc điều tra.

2.6. Xây dựng bảng hỏi trong điều tra thống kê

2.6.1. Bảng hỏi và yêu cầu của việc xây dựng bảng hỏi trong điều tra thống kê

Bảng hỏi (hay phiếu điều tra) là hệ thống các câu hỏi được sắp xếp trên cơ sở các nguyên tắc, trình tự logic và theo nội dung nhất định nhằm giúp cho người điều tra có thể thu thập thông tin về hiện tượng nghiên cứu một cách đầy đủ, đáp ứng mục tiêu nghiên cứu đã đề ra.

Trong điều tra thống kê, bảng hỏi là một công cụ quan trọng giúp truyền tải yêu cầu thông tin từ người điều tra đến đối tượng được điều tra (câu hỏi) và thu nhận thông tin ngược lại (câu trả lời). Việc đo lường, nhận thức hiện tượng có đầy đủ hay không phụ thuộc chặt chẽ vào chỗ các câu hỏi này có bao quát được nội dung nghiên cứu hay không. Chính vì vậy, khi nhìn vào bảng hỏi, ta có thể biết cuộc điều tra này được tiến hành nhằm giải quyết vấn đề gì.

Bảng hỏi là một công cụ đo lường quan trọng, nhờ đó ta có thể đo được các biến số nhất định có liên quan đến chủ đề cần nghiên cứu. Vì vậy, nếu bảng hỏi được soạn thảo tốt sẽ cho ta những thông tin đầy đủ, khoa học, ngược lại sẽ thu được những thông tin sai lệch.

Bảng hỏi còn giữ vai trò cầu nối giữa nhân viên điều tra với người trả lời câu hỏi. Nó chịu ảnh hưởng của người nghiên cứu khi muốn chuyển các nội dung cần thiết thành các câu hỏi cụ thể với mong muốn thu được những thông tin chính xác. Mặt khác, các câu hỏi trong bảng hỏi cũng phải phù hợp với trình độ của người được hỏi. Muốn vậy, bảng hỏi phải được thiết kế đảm bảo yêu cầu mỹ thuật, tiết kiệm, tiện dụng.

2.6.2. Các loại câu hỏi và kỹ thuật đặt các loại câu hỏi

2.6.2.1. Các loại câu hỏi

- ***Câu hỏi theo nội dung***

Cơ sở phân loại câu hỏi theo nội dung là thực tế kinh tế - xã hội mà các câu hỏi này đề cập đến và có thể truyền tải được. Theo nội dung, ta có thể chia chi tiết thành các câu hỏi về kinh tế, văn hoá, xã hội... Về kinh tế có thể phân biệt câu hỏi theo từng ngành: Công nghiệp, nông nghiệp... Tuy nhiên, theo cách hiểu về kỹ thuật xây dựng bảng hỏi, ta có thể phân biệt hai nhóm câu hỏi theo nội dung như sau:

- + ***Nhóm thứ nhất: Câu hỏi về sự kiện***

Là những câu hỏi nhằm thu thập những thông tin thực tế gắn với đối tượng điều tra (thông tin về thân thế, sự nghiệp, lứa tuổi, giới tính, địa vị xã hội...) và những sự kiện đã xảy ra đối với đối tượng điều tra. Những câu hỏi này được đặt ra nhằm để nắm tình hình hiện thực khách quan, bao gồm cả tình hình về đối tượng điều tra.

Ví dụ: Năm vừa qua doanh số bán hàng của công ty là bao nhiêu? Thu nhập trung bình của một lao động hiện tại là bao nhiêu? Hoặc những câu hỏi về giới tính, trình độ văn hoá, nghề nghiệp...

Loại câu hỏi này thường có mặt trong hầu hết các phiếu điều tra, các bảng hỏi, ngoại trừ các câu hỏi riêng tư. Câu hỏi về các sự kiện nói chung đều dễ trả lời. Vì vậy, loại câu hỏi này thường hay sử dụng để làm quen với người điều tra.

+ Nhóm thứ hai: Câu hỏi tri thức (câu hỏi hiểu biết)

Là loại câu hỏi nhằm đánh giá trình độ hiểu biết của người được hỏi về một vấn đề. Tuy nhiên, tính chuẩn xác của câu trả lời tương ứng sẽ phụ thuộc vào nhận thức của người được hỏi.

Ví dụ: Bạn hiểu gì về chuyên ngành đào tạo mà bạn đang theo học? Trình độ tiếng Anh của bạn ở mức nào?

Việc phân chia 2 loại câu hỏi như trên có liên quan đến phương pháp thu thập thông tin. *Đối với nhóm thứ nhất*, các sự kiện được hỏi luôn tồn tại và gắn liền với hiện thực cuộc sống, chúng có tính khách quan nên dễ trả lời, thông tin thu được cũng có độ chính xác cao. *Đối với nhóm thứ hai*, nó liên quan đến sự đánh giá, nhìn nhận vấn đề của từng cá nhân nên khó khăn trong việc diễn đạt. Hơn nữa, nhiều đánh giá, nhận định liên quan đến vấn đề riêng tư làm cho việc trả lời gặp khó khăn, thậm chí có trường hợp không thể thu được thông tin chính xác. Vì vậy, khi sử dụng nhóm thứ hai cần quan tâm đến cách đặt câu hỏi, cách sử dụng từ ngữ, hình thức câu hỏi.

• *Câu hỏi chức năng*

Trên thực tế, để chuyển tải những nội dung của thông tin trong điều tra, đặc biệt là khi sử dụng phương pháp phỏng vấn cần phải có những câu hỏi mang tính chất kỹ thuật, đó là các câu hỏi chức năng bao gồm câu hỏi tâm lý, câu hỏi lọc, câu hỏi kiểm tra.

- *Câu hỏi tâm lý*: Có thể là câu hỏi tiếp xúc nhằm mục đích gạt bỏ những nghi ngờ có thể nảy sinh, để giảm bớt sự căng thẳng, hoặc chuyển từ chủ đề này sang chủ đề khác... thường dùng trong phỏng vấn trực tiếp.

Để tiếp xúc, làm quen, người ta thường đưa ra những câu hỏi rất đơn giản, dễ trả lời, thậm chí có thể đã biết câu trả lời. Mục đích của việc đặt câu hỏi này không nhất thiết là để thu thập thông tin mà chủ yếu mang tính chất làm quen.

Để giảm bớt sự căng thẳng, có thể biểu thị sự quan tâm tới người được hỏi, người ta thường đặt chen vào giữa những câu hỏi nội dung một số câu hỏi về đời sống hằng ngày, về gia đình... sau đó mới trở về vấn đề cần hỏi.

- Câu hỏi lọc:

Để chuyển từ nội dung này sang nội dung khác, có thể sử dụng câu hỏi chuyển tiếp để làm người được hỏi không cảm thấy hụt hẫng, nghi ngờ, thắc mắc tính hợp lý của câu hỏi. Loại câu hỏi này có tác dụng tìm hiểu xem người được hỏi có thuộc nhóm người dành cho những câu hỏi tiếp sau không.

Khi xây dựng bảng hỏi, người ta thường sử dụng kỹ thuật “bước nhảy” với ý nghĩa là cho phép chuyển đến câu hỏi khác.

Ví dụ: Trong phiếu điều tra Tổng dân số năm 1999, có câu hỏi:

Câu 6. Ông (bà) có theo tôn giáo (đạo) nào không?

1. Có ☐ 2. Không ☐ Nếu “có” Xin chuyển sang hỏi câu 9

- Câu hỏi kiểm tra:

Có tác dụng kiểm tra tính chính xác của thông tin thu được. Nó được sử dụng khi gặp một câu trả lời bị nghi ngờ về tính xác thực. Thông thường người ta gài thêm câu hỏi mà đã biết chắc chắn câu trả lời để kiểm tra tính trung thực của người trả lời.

Một số vấn đề cần chú ý khi sử dụng câu hỏi kiểm tra:

+ Các câu hỏi kiểm tra thường làm tăng số câu hỏi trong bảng hỏi, nếu sử dụng quá nhiều sẽ ảnh hưởng đến chất lượng cuộc điều tra.

+ Câu hỏi kiểm tra không bao giờ được đi liền với câu trả lời mà nó kiểm tra và thường là đặt cách xa khoảng 3 đến 4 câu.

+ Sử dụng câu hỏi kiểm tra cũng phải hết sức khéo léo, cần tránh cho người trả lời có cảm giác tự ái, cảm thấy bị xúc phạm.

- *Câu hỏi theo cách biểu hiện*

Đây là cách phân loại câu hỏi theo hình thức biểu hiện của câu trả lời cũng như cách hỏi.

- *Theo biểu hiện của câu trả lời:*

Câu hỏi đóng: Là dạng câu hỏi có trước phương án trả lời cụ thể và người trả lời chỉ việc chọn một trong số các phương án trên. Loại câu hỏi này thường dễ trả lời và thuận tiện trong quá trình tổng hợp.

Câu hỏi mở: Là câu hỏi không có phương án trả lời được nêu trước mà nó hoàn toàn do người trả lời tự nghĩ ra. Chúng cho phép người được hỏi tự cung cấp thông tin những suy nghĩ của họ.

- *Theo cách biểu hiện của câu hỏi:*

Câu hỏi trực tiếp: Là cách hỏi thẳng ngay vào nội dung vấn đề, người được hỏi không bị quanh co và có thể trả lời vào chính nội dung đó.

Câu hỏi gián tiếp: Là cách hỏi khôn khéo, không đi trực tiếp vào vấn đề, mà thông qua những vấn đề có liên quan để thu thập thông tin về vấn đề cần hỏi. Câu hỏi này thường dùng cho những vấn đề mà xã hội gán cho nó tính tiêu cực, những chủ đề riêng tư, sâu kín trọng tình cảm.

2.6.2.2. *Kỹ thuật đặt câu hỏi*

Việc đặt câu hỏi sao cho đối tượng trả lời nắm được đầy đủ ý nghĩa, hiểu được nội dung và có thông tin để trả lời là những vấn đề có ý nghĩa quan trọng trong phỏng vấn thu thập thông tin. Cùng một nội dung nhưng có người đặt những câu hỏi được đánh giá là “hay”, nhưng có người lại đặt những câu hỏi mà đối tượng trả lời không hiểu được nội dung để trả lời hoặc trả lời không đúng nội dung mà người đặt câu hỏi cần hỏi. Vì vậy, xây dựng các câu hỏi một cách hiệu quả đòi hỏi người xây dựng phải có trình độ chuyên môn, kinh nghiệm thực tế, kiến thức cơ bản về lĩnh vực nghiên cứu, ngôn ngữ và hiểu biết tâm lý đóng vai trò quan trọng.

Người thiết kế câu hỏi thường mắc phải sai lầm phổ biến khi nghĩ rằng các đối tượng điều tra sẽ hiểu ngay câu hỏi và đưa ra câu trả lời có

cùng một nghĩa của thuật ngữ và khái niệm giống như họ. Nhưng thực tế lại không như vậy, đối tượng điều tra có thể không nghĩ về các câu hỏi một cách chi tiết như những người thiết kế yêu cầu. Cho nên, khi đặt một câu hỏi, trước hết người thiết kế câu hỏi phải tự đặt mình vào vị trí một đối tượng điều tra nhất định, hoặc một đối tượng ít được đào tạo nhất. Thực tế cho thấy, những cuộc điều tra không mang lại nhiều kết quả, thông tin thu được không như mong muốn thường bắt nguồn từ việc soạn thảo, chuẩn bị và đặt câu hỏi không đúng nguyên tắc. Mặc dù không có cách thức chính thức về cách đặt một câu hỏi nhưng việc đặt các câu hỏi vẫn phải tuân theo những nguyên tắc chung nhất định. Khi đặt câu hỏi, người thiết kế cần phải chú ý đến một số vấn đề sau:

- Lựa chọn câu hỏi đúng loại và phù hợp, để làm được điều này cần nắm được đặc điểm và nguyên tắc của từng loại câu hỏi. Khi đặt câu hỏi người ta thường nghĩ đến mục đích của việc đặt câu hỏi là có được thông tin. Tuy nhiên, trong thiết kế bảng hỏi, mỗi câu hỏi có những đặc điểm, tác dụng và nguyên tắc xây dựng khác nhau. Người thiết kế cần nắm được những đặc điểm và nguyên tắc này để đặt những câu hỏi cho phù hợp.

- Sắp xếp các câu hỏi hợp lý, logic, việc đặt câu hỏi theo một trật tự nào đó có ý nghĩa quan trọng đối với tính khách quan của những thông tin thu thập được. Theo đó, các câu hỏi cần được bố trí theo các hướng sau:

- + Các câu hỏi cần đặt theo chủ đề, những câu hỏi có cùng chủ đề cần được đặt liên nhau nhằm đảm bảo tính hệ thống, xuyên suốt.

- + Các câu hỏi được sắp xếp theo trật tự nhất định từ vấn đề đơn giản đến vấn đề phức tạp, từ vấn đề nhận thức được đến vấn đề không nhận thức được, từ cái chung đến cái riêng, từ câu hỏi mang những nội dung chung, tổng quát đến những câu hỏi có những nội dung cụ thể hơn.

- + Đặt các câu hỏi dễ trả lời trước, thường đặt những câu hỏi đầu tiên là những câu hỏi tâm lý, câu hỏi sự kiện nhằm tạo tâm lý thoải mái và thúc đẩy tinh thần tham gia phỏng vấn cho người được hỏi. Không nên đặt câu hỏi về thái độ, động cơ hay đánh giá về vấn đề nào đó khi mới bước vào phỏng vấn.

+ Đặt câu hỏi nhạy cảm, câu hỏi mở, câu hỏi khó ở cuối cùng để tránh cho đối tượng “bị kẹt” ngay ở những câu hỏi đầu.

+ Cần thay đổi độ dài của các câu hỏi, sử dụng đa dạng các loại câu hỏi, các thang điểm tránh nhàm chán và dẫn đến trả lời theo quán tính (các câu trả lời tương tự nhau) và chú ý logic về nội dung thông tin cũng như tiến trình thời gian hợp lý của các câu hỏi.

- Một trình tự câu hỏi hợp lý, theo logic suy nghĩ của đối tượng trả lời sẽ tạo ra tâm lý hứng thú trong việc tham gia trả lời các câu hỏi. Đồng thời cũng cần chú ý đến cách hướng dẫn trả lời cần ngắn gọn, dễ hiểu, càng gần câu hỏi càng tốt, hạn chế những hướng dẫn dài dòng.

- Một yếu tố cũng rất quan trọng và cần được quan tâm khi đặt câu hỏi đó là yếu tố tâm lý. Chú ý đến khía cạnh tâm lý khi đặt câu hỏi nhằm tránh gây ra những mệt mỏi, căng thẳng đối với đối tượng trả lời. Thông thường, sau những câu hỏi về quan điểm, thái độ, động cơ mang tính nhận xét, đánh giá về một vấn đề nào đó thường là những câu hỏi tâm lý để chuyển sang một chủ đề khác.

2.7. Sai số trong điều tra thống kê

Bất kỳ một cuộc điều tra thống kê nào dù được tổ chức rất tốt nhưng vẫn thường mắc phải những sai lầm dẫn đến những sai số nhất định. Các sai số này làm giảm chất lượng của kết quả điều tra và có ảnh hưởng tới kết quả tổng hợp, phân tích thống kê. Bởi vậy, vấn đề đặt ra là phải nắm được các nguyên nhân phát sinh sai số trong điều tra thống kê, để chủ động đề ra các biện pháp khắc phục hoặc hạn chế sai số đó.

2.7.1. Khái niệm

Sai số trong điều tra thống kê là chênh lệch giữa trị số thực của hiện tượng nghiên cứu so với trị số của nó mà điều tra thống kê thu được.

Có thể phân biệt 2 loại sai số:

+ *Sai số do đăng ký*: Là sai số phát sinh do việc đăng ký số liệu ban đầu không chính xác. Nguyên nhân của loại sai số này có rất nhiều: có thể do điều tra viên vô tình cân, đong, đo, đếm, ghi chép sai; hoặc do đơn

vị điều tra chưa hiểu hết ý nghĩa câu hỏi nên trả lời sai; có khi do nhân viên điều tra cố tình ghi sai... sai số này có thể khắc phục được.

+ *Sai số do tính chất đại diện* là sai số gặp phải trong điều tra chọn mẫu. Nguyên nhân là do việc lựa chọn số đơn vị điều tra không đủ tính chất đại biểu: Đặc điểm của mẫu không phù hợp đặc điểm của cả tổng thể, nên khi suy rộng thì mắc phải sai số này.

2.7.2. Biện pháp hạn chế sai số

Để đảm bảo các kết quả điều tra đạt mức độ chính xác cao, cần tìm cách khắc phục hoặc hạn chế bớt một phần sai số nói trên. Các biện pháp để hạn chế sai số bao gồm:

+ *Làm tốt công tác chuẩn bị điều tra, cụ thể:* Xây dựng phương án điều tra khoa học và theo đúng nội dung, mục đích điều tra; Chuẩn bị cán bộ, lựa chọn, tập huấn, giáo dục tư tưởng; Mặt khác cần coi trọng công tác tuyên truyền mục đích và ý nghĩa của cuộc điều tra trong quần chúng nhân dân nhất là các cuộc điều tra lớn.

+ *Kiểm tra một cách có hệ thống toàn bộ cuộc điều tra:* Kiểm tra tài liệu thu thập được xem có đầy đủ hay không về nội dung và số đơn vị điều tra, kiểm tra tính chính xác của tài liệu thu thập được về con số và về mặt logic. Ở bước này việc nghiệm thu phiếu điều tra có ý nghĩa quan trọng, cụ thể cần phải xem xét các dữ liệu trong phần trả lời có hợp lý, logic hay không. Công việc này cần phải có sự tham gia của các chuyên gia có trình độ chuyên môn về thống kê và có hiểu biết rộng về kiến thức thực tế. Trường hợp phát hiện những phiếu điều tra không đảm bảo thì cần phải loại bỏ và tiến hành điều tra bổ sung. Kiểm tra tính chất đại biểu của số đơn vị được chọn trong điều tra chọn mẫu, có thể so sánh một số đặc điểm cơ bản của số đơn vị được chọn với các đặc điểm tương ứng của tổng thể chung, kết cấu mẫu so với kết cấu tổng thể chung.

+ *Phục tra lại kết quả điều tra:* Công việc này được tiến hành sau khi hoàn thành quá trình điều tra. Phục tra là việc thu thập lại thông tin với các đối tượng đã được điều tra nhằm đánh giá mức độ chính xác và làm cơ sở để có thể chỉnh lý lại số liệu đã có được. Tuy nhiên, để giảm chi phí điều tra, việc phục tra thường tiến hành theo phương pháp chọn mẫu và tập trung vào những nơi nghi ngờ thường hay bị khai báo sai.

+ *Kiểm tra quá trình nhập số liệu vào máy tính*: Thực tế cho thấy đây cũng là khâu dễ làm phát sinh sai số. Có nhiều trường hợp phải tiến hành nhập số liệu hai lần độc lập để khắc phục sai số trong quá trình nhập dữ liệu.

Kiểm tra là một biện pháp có hiệu quả để kịp thời sửa chữa các sai lầm có thể gặp phải trong điều tra thống kê.

TÓM TẮT CHƯƠNG 2

Điều tra thống kê là việc tổ chức một cách khoa học theo một kế hoạch thống nhất việc thu thập, ghi chép nguồn tài liệu ban đầu về hiện tượng nghiên cứu trong điều kiện cụ thể về thời gian, không gian. Khi tiến hành điều tra thống kê phải đảm bảo ba yêu cầu: Chính xác, kịp thời và đầy đủ. Đây là những yêu cầu quan trọng đảm bảo tính khách quan, hữu dụng của các thông tin thu thập được. Căn cứ vào tính chất liên tục hay không liên tục có thể chia các cuộc điều tra thống kê được thành hai loại: Điều tra thường xuyên và điều tra không thường xuyên. Căn cứ vào phạm vi của đối tượng được điều tra thực tế, điều tra thống kê được chia thành điều tra toàn bộ và điều tra không toàn bộ. Khi tiến hành thu thập tài liệu ban đầu có thể thu thập bằng phương pháp đăng ký trực tiếp hoặc phương pháp phỏng vấn. Tùy thuộc vào đặc điểm của đối tượng nghiên cứu và điều kiện thực tế để lựa chọn phương pháp thu thập tài liệu phù hợp. Để tổ chức một cuộc điều tra cần thiết phải xây dựng một phương án điều tra thích hợp. Phương án điều tra thường bao gồm những nội dung chủ yếu sau:

- Xác định mục đích điều tra.
- Xác định phạm vi đối tượng và đơn vị điều tra.
- Xác định nội dung và phiếu điều tra.
- Lựa chọn thời điểm, thời kỳ và xác định thời hạn điều tra.
- Các danh mục và bảng phân loại.
- Loại điều tra và phương pháp thu thập thông tin.
- Lập kế hoạch tổ chức và tiến hành điều tra.

Trong đó việc thiết kế phiếu điều tra (bảng hỏi) là quan trọng nhất, vì đây là công cụ truyền tải toàn bộ thông tin từ điều tra viên đến đối tượng nghiên cứu và thu nhận thông tin ngược lại.

Do tính chất phức tạp của đối tượng điều tra, trong điều tra thống kê thường gặp phải sai số điều tra. Sai số trong điều tra thống kê có thể phân biệt là sai số do đăng ký và sai số do tính chất đại diện (sai số chọn mẫu). Sai số làm giảm chất lượng của tài liệu điều tra. Để hạn chế sai số phải sử dụng đồng bộ các biện pháp: Làm tốt công tác chuẩn bị điều tra, tiến hành kiểm tra có hệ thống toàn bộ cuộc điều tra.

CÂU HỎI ÔN TẬP CHƯƠNG 2

1. Trình bày khái niệm, ý nghĩa và yêu cầu cơ bản của điều tra thống kê?
2. Trình bày các loại điều tra thống kê?
3. Trình bày các phương pháp thu thập thông tin trong điều tra thống kê?
4. Trình bày các hình thức tổ chức điều tra thống kê?
5. Trình bày những vấn đề cơ bản trong xây dựng phương án điều tra?
6. Trình bày khái niệm, phân loại và các biện pháp khắc phục sai số trong điều tra thống kê?

BÀI TẬP CHƯƠNG 2

Bài 1:

Hãy xác định và giải thích các cuộc điều tra thống kê sau đây thuộc về hình thức tổ chức, loại và phương pháp điều tra gì?

1. Tổng điều tra dân số và nhà ở toàn quốc vào ngày 1/4/2009.
2. Các cuộc tổng điều tra thống kê về năng suất và sản lượng lúa, hoa màu và cây công nghiệp.
3. Điều tra tình hình giá cả một số loại hàng hóa trên thị trường.

4. Báo cáo tình hình hàng hoá tồn kho (0 giờ ngày 1/1 và 1/7 hàng năm) của các doanh nghiệp thương mại.

5. Báo cáo tình hình bán ra (tháng, quý, năm) của các công ty thương mại.

6. Báo cáo hàng tháng về tình hình hoàn thành kế hoạch giá trị tổng sản lượng và sản lượng hàng hoá của các xí nghiệp công nghiệp.

Bài 2:

Địa phương X cần tổ chức một cuộc điều tra thống kê nhằm nghiên cứu số lượng, kết cấu và trình độ văn hoá của các cán bộ, giáo viên ngành giáo dục phổ thông trong địa phương. Yêu cầu xác định:

1. Đối tượng điều tra?
2. Đơn vị điều tra?
3. Nội dung điều tra?
4. Biểu điều tra và bản giải thích cách ghi biểu?
5. Phương pháp điều tra?

Bài 3:

Hãy điền đúng (Đ) hoặc sai (S) vào các mục sau:

1. Hệ thống chỉ tiêu thống kê phản ánh:
 - a) Các đặc điểm quan trọng nhất của tổng thể.
 - b) Các mặt quan trọng của đơn vị tổng thể.
 - c) Kết cấu tổng thể.
 - d) Các mặt, các tính chất quan trọng của tổng thể.
 - e) Mối liên hệ của tổng thể với các hiện tượng khác.
2. Điều tra thống kê là:
 - a) Kiểm tra kết quả tài liệu thu thập được.
 - b) Bước cuối cùng của quá trình nghiên cứu thống kê.
 - c) Năm đặc điểm, tình hình của hiện tượng.
 - d) Ghi chép, thu thập tài liệu ban đầu của hiện tượng.
 - e) Tính toán, so sánh mặt lượng của hiện tượng.

3. Thời kì điều tra là:

- a) Độ dài thời gian cuộc điều tra.
- b) Bắt đầu cuộc điều tra thống kê.
- c) Thời gian nhân viên đi thu thập tài liệu.
- d) Thời gian chuẩn bị xây dựng phương án điều tra.
- e) Là khoảng thời gian được quy định để thu thập số liệu về hiện tượng được tích lũy trong cả thời kỳ đó.

Bài 4:

Hãy trả lời đúng (hoặc sai) cho các câu hỏi sau đây:

1. Tổng hợp thống kê là:

- a) Tập hợp tất cả các tài liệu của hiện tượng nghiên cứu.
- b) Tổng hợp thống kê là giai đoạn cuối cùng của quá trình nghiên cứu thống kê.
- c) Tổng hợp thống kê là tiến hành phân chia các đơn vị của tổng thể.
- d) Tổng hợp thống kê là sự tập trung, chỉnh lý và hệ thống hoá tài liệu đã thu thập được của tổng thể.
- e) Tổng hợp thống kê là đánh giá, nhận xét chung về hiện tượng nghiên cứu.

2. Phân tích thống kê là:

- a) Sự tập trung, hệ thống hoá tài liệu thu thập được qua điều tra.
- b) Xử lý tài liệu thu thập được.
- c) Toàn bộ quá trình nghiên cứu thống kê.
- d) Tính toán mức độ của hiện tượng trong tương lai.
- e) Biểu hiện bản chất cụ thể và tính quy luật của hiện tượng thông qua mặt lượng.

Chương 3

TỔNG HỢP THỐNG KÊ

Với đặc điểm của khoa học thống kê là nghiên cứu số lớn, do vậy những tài liệu thu thập được trên đơn vị trong điều tra thống kê chưa sử dụng trực tiếp cho phân tích và dự đoán. Do vậy, cần tiến hành sắp xếp, chỉnh lý, hệ thống hóa tài liệu thu thập được để bước đầu thể hiện được những đặc trưng chung của tổng thể nghiên cứu, đây là giai đoạn tổng hợp thống kê. Nội dung chương này sẽ trình bày những vấn đề chung của tổng hợp thống kê, phương pháp cơ bản để tiến hành tổng hợp là phương pháp phân tổ thống kê và trình bày tài liệu tổng hợp bằng bảng thống kê và đồ thị thống kê.

3.1. Những vấn đề chung của tổng hợp thống kê

3.1.1. *Khái niệm, ý nghĩa và nhiệm vụ của tổng hợp thống kê*

Sau khi tiến hành điều tra thống kê, ta thu được những tài liệu về tiêu thức điều tra trên mỗi đơn vị điều tra. Các tài liệu này mới chỉ phản ánh các đặc trưng riêng của từng đơn vị tổng thể. Để có thể nêu lên một số đặc trưng chung của tổng thể, người ta phải tổng hợp các tài liệu điều tra.

Tổng hợp thống kê là tiến hành tập trung chỉnh lý và hệ thống hóa một cách khoa học các tài liệu thu thập được trong điều tra thống kê.

Ví dụ, Tài liệu điều tra dân số vào 0 giờ ngày 01 tháng 4 năm 2009 của nước ta được tổng hợp cho biết một số đặc điểm cơ bản của dân số Việt Nam vào thời điểm trên.

Nhiệm vụ cơ bản của tổng hợp thống kê là làm cho các đặc trưng riêng của đơn vị tổng thể bước đầu chuyển thành các đặc trưng chung của tổng thể.

Tổng hợp thống kê là giai đoạn thứ hai của quá trình nghiên cứu thống kê. Việc tổ chức tổng hợp một cách khoa học sẽ có ý nghĩa rất lớn đối với kết quả nghiên cứu thống kê, bởi các đặc trưng chung của tổng

thể là căn cứ cho phân tích và dự báo thống kê. Tổng hợp thống kê đúng đắn làm cho kết quả điều tra trở nên có giá trị và tạo điều kiện cho phân tích sâu sắc bản chất, tính quy luật phát triển của hiện tượng.

Tổng hợp thống kê là công việc phức tạp, phải xây dựng kế hoạch thống nhất trong đó quy định những vấn đề cần giải quyết và cần được quan tâm thích đáng trong quá trình tổng hợp.

3.1.2. Các vấn đề chủ yếu của tổng hợp thống kê

3.1.2.1. Mục đích của tổng hợp thống kê

Phân tích và dự báo thống kê phải dựa vào những tài liệu biểu hiện hình ảnh thực của tổng thể nghiên cứu. Kết quả tổng hợp là căn cứ cho phân tích và dự báo. Cho nên mục đích của tổng hợp thống kê là khái quát hóa những đặc trưng chung, những cơ cấu tồn tại khách quan theo các mặt của tổng thể nghiên cứu bằng các chỉ tiêu thống kê. Khi xác định mục tiêu của tổng hợp thống kê phải căn cứ vào mục đích, yêu cầu tìm hiểu và phân tích những mặt nào của hiện tượng nghiên cứu, để nêu khái quát những chỉ tiêu cần đạt được trong tổng hợp.

3.1.2.2. Nội dung tổng hợp

Nội dung tổng hợp được xác định để đáp ứng mục đích tổng hợp. Nội dung tổng hợp là danh mục các biểu hiện của những tiêu thức mà chúng được xác định trong nội dung điều tra. Tuy nhiên, không phải tất cả các biểu hiện của tiêu thức điều tra đều đưa vào nội dung tổng hợp, mà phải chọn lọc để nội dung tổng hợp vừa đủ đáp ứng mục đích nghiên cứu. Nói một cách cụ thể, nội dung tổng hợp là danh mục các biểu hiện của tiêu thức điều tra được chọn lọc và theo mỗi biểu hiện được phân chia thành các nhóm khác nhau để đáp ứng yêu cầu phản ánh các cơ cấu khác nhau phù hợp với mục đích nghiên cứu. Nội dung tổng hợp cũng là danh mục của một hệ thống chỉ tiêu tổng hợp. Ví dụ: Một trong những nội dung điều tra dân số là tuổi của nhân khẩu, biểu hiện cụ thể của nó là các độ tuổi khác nhau. Người ta phải tổng hợp theo từng độ tuổi trên để biết từng độ tuổi có bao nhiêu người. Từng độ tuổi lại được phân chia theo giới tính (nam và nữ) theo trình độ văn hóa (từ 6 tuổi trở lên), theo nơi cư trú (thành thị và nông thôn), theo dân tộc...

3.1.2.3. Kiểm tra tài liệu dùng vào tổng hợp

Kiểm tra tài liệu dùng vào tổng hợp là một việc làm không thể thiếu vì chất lượng kết quả tổng hợp phụ thuộc vào chất lượng tài liệu dùng vào tổng hợp.

Kiểm tra tài liệu đã được thực hiện trong khâu điều tra. Việc kiểm tra này được tiến hành trên nhiều mặt, kiểm tra toàn bộ, do những người trực tiếp tham gia điều tra làm. Tuy vậy; khi tổng hợp vẫn phải kiểm tra tài liệu một lần nữa, để đảm bảo tính chính xác của tài liệu, loại bỏ hẳn tất cả hay một phần nội dung của những phiếu điều tra không đúng, nếu không có điều kiện điều tra lại.

Đối với các cuộc điều tra lớn, khối lượng phiếu điều tra nhiều, không thể kiểm tra toàn bộ được, người ta chọn mẫu một số phiếu điều tra để kiểm tra.

Khi kiểm tra phải xét xem các tài liệu thu thập được có chính xác không (nội dung kiểm tra cụ thể như trong khâu điều tra).

3.1.2.4. Phương pháp tổng hợp

Yêu cầu quan trọng nhất của tổng hợp là phải nêu lên được cơ cấu theo các mặt của tổng thể nghiên cứu. Để đáp ứng yêu cầu này người ta sử dụng phương pháp phân tổ thống kê.

Phân tổ thống kê là phương pháp phân chia các đơn vị của tổng thể vào các tổ (và tiểu tổ) khác nhau theo từng tiêu thức nghiên cứu. Trị số của mỗi tổ được tính toán sẽ cho một cơ cấu về lượng cụ thể của tổng thể. Phương pháp phân tổ thống kê có nội dung phong phú sẽ được trình bày trong phần tiếp theo.

3.1.2.5. Tổ chức và kỹ thuật tổng hợp thống kê

Tổ chức và kỹ thuật tổng hợp thống kê không chỉ đơn thuần là lĩnh vực của thống kê mà là lĩnh vực của cả tin học. Cho nên, trong thực tế, ở đây luôn có sự kết hợp chặt chẽ giữa thống kê và tin học.

*** Chuẩn bị tài liệu để tổng hợp.**

Mỗi cơ quan hay cá nhân thực hiện tổng hợp đều phải tập trung đầy đủ số lượng phiếu điều tra hoặc tài liệu khác để có thể đáp ứng việc thực

hiện nhiệm vụ được đảm nhiệm. Tài liệu tập trung không đầy đủ từ đầu mà tiến hành tổng hợp, sau đó phải tiến hành tổng hợp bổ sung sẽ làm cho khối lượng công việc tổng hợp tăng thêm gần như một lần, mỗi khi tổng hợp bổ sung.

*** Hình thức tổ chức tổng hợp.**

Có hai hình thức tổ chức tổng hợp thống kê chủ yếu: Từng cấp và tập trung.

Tổng hợp từng cấp là tổ chức tổng hợp tài liệu điều tra theo từng bước, từ cấp dưới lên cấp trên theo kế hoạch đã vạch sẵn. Cơ quan phụ trách tổng hợp các cấp tiến hành tổng hợp tài liệu trong phạm vi được phân công, sau đó gửi kết quả lên cấp cao hơn để tiến hành tổng hợp theo phạm vi rộng hơn. Theo trình tự như vậy, cuối cùng các tài liệu được gửi về cấp trên, ở đây sẽ tiến hành tổng hợp lần cuối, tính ra các chỉ tiêu chung nêu rõ tình hình của toàn bộ hiện tượng nghiên cứu.

Tổng hợp tập trung là toàn bộ tài liệu điều tra được tập trung về một cơ quan để tiến hành tổng hợp từ đầu đến cuối.

Mỗi hình thức tổ chức tổng hợp trên đều có một số ưu điểm và hạn chế nhất định. Tổng hợp từng cấp thường áp dụng trong báo cáo thống kê định kỳ và một số cuộc điều tra chuyên môn. Ở nước ta hiện nay, hình thức tổ chức tổng hợp này là chủ yếu. Số liệu do từng cấp tổng hợp đưa lên thường có mức độ chính xác cao, vì các cán bộ tổng hợp mỗi cấp nói chung đều hiểu biết tình hình địa phương hoặc đơn vị mình, có thể tự kiểm tra phát hiện sai sót trong tài liệu điều tra. Mặt khác tổng hợp từng cấp tương đối gọn, phục vụ kịp thời cho yêu cầu về thông tin của từng cấp. Tuy nhiên, tổng hợp từng cấp tốn nhiều công sức mà lại chậm. Mặt khác, phạm vi tổng hợp từng cấp thường nhỏ, kết quả tổng hợp chỉ gồm một số chỉ tiêu nhất định, hạn chế. Tổng hợp tập trung thường chỉ được áp dụng đối với một số cuộc điều tra lớn như tổng điều tra dân số. Trong tổng hợp tập trung, thường người ta phải sử dụng những phương tiện hiện đại để tính toán nhanh chóng và chính xác những chỉ tiêu phức tạp. Do đó, tổng hợp tập trung giảm bớt được nhiều công việc thủ công vất vả. Nhưng, việc cung cấp kết quả tổng hợp để phục vụ cho các cấp dưới thường không nhanh.

Trong thực tế người ta có thể kết hợp hai hình thức tổ chức tổng hợp với nhau. Căn cứ vào yêu cầu nghiên cứu của các cấp, cơ quan tổng hợp trung ương giao cho các cấp tổng hợp một số chỉ tiêu đáp ứng yêu cầu nghiên cứu trước mắt. Sau đó gửi kết quả và toàn bộ tài liệu về cơ quan tổng hợp trung ương để tổng hợp theo kế hoạch đã định. Với hình thức tổng hợp này, kết quả tổng hợp phục vụ kịp thời cho các cấp dưới và trung ương.

3.2. Phân tổ thống kê

3.2.1. Khái niệm, ý nghĩa, nhiệm vụ của phân tổ thống kê

Phân tổ thống kê là căn cứ vào một (hay một số) tiêu thức nào đó tiến hành phân chia các đơn vị của hiện tượng nghiên cứu thành các tổ (và các tiểu tổ) có tính chất khác nhau.

Chẳng hạn, khi nghiên cứu tình hình nhân khẩu nếu căn cứ vào tiêu thức giới tính thì có thể chia tổng số nhân khẩu thành hai tổ: Nam, nữ hoặc nếu căn cứ theo tiêu thức tuổi thì có thể chia nhân khẩu ra các nhóm có độ tuổi khác nhau. Khi nghiên cứu tình hình sản xuất của các doanh nghiệp, có thể chia tổng số doanh nghiệp thành các nhóm theo các tiêu thức như: Thành phần kinh tế, ngành sản xuất, số lượng công nhân, giá trị sản xuất v.v...

Khi phân tổ thống kê, các đơn vị tổng thể được tập hợp lại thành một số tổ, giữa các tổ có sự khác nhau về tính chất, còn trong phạm vi mỗi tổ các đơn vị đều có sự giống nhau (hoặc gần giống nhau) về tính chất theo tiêu thức được dùng làm căn cứ phân tổ.

Phân tổ thống kê có nhiều ý nghĩa trong nghiên cứu thống kê. Trong nhiều trường hợp khi tiến hành điều tra người ta đã dùng đến phân tổ. Chẳng hạn như điều tra doanh thu của những người kinh doanh không thể điều tra tất cả, trước hết phải chia số người kinh doanh theo ngành hàng, nhóm hàng kinh doanh để thu thập số liệu của một số người trong ngành, nhóm hàng.

Phân tổ là phương pháp cơ bản để tiến hành tổng hợp thống kê. Muốn hệ thống hóa một cách khoa học các tài liệu điều tra, muốn tổng

hợp theo những chỉ tiêu đã nêu ra, thì phải căn cứ vào từng chỉ tiêu đó mà sắp xếp các đơn vị vào từng tổ, và sau đó mới tính đặc trưng chung của cả tổng thể.

Chẳng hạn, muốn tổng hợp chỉ tiêu trình độ văn hóa sau khi điều tra dân số thì phải sắp xếp số người có trình độ văn hóa khác nhau (phổ thông cơ sở, phổ thông trung học, đại học...) theo các nhóm khác nhau, nghĩa là phải phân tổ nhân khẩu theo tiêu thức trình độ văn hóa.

Phân tổ thống kê là một trong các phương pháp quan trọng của phân tích thống kê, đồng thời là cơ sở để vận dụng các phương pháp thống kê khác. Phương pháp phân tổ được vận dụng phổ biến nhất trong mọi trường hợp nghiên cứu kinh tế - xã hội, vì phương pháp này đơn giản, dễ hiểu và có tác dụng phân tích sâu sắc.

Hiện tượng và quá trình kinh tế - xã hội mà thống kê học nghiên cứu thường phức tạp. Vì chúng tồn tại và phát triển dưới nhiều loại hình có quy mô và đặc điểm khác nhau. Trong bản thân loại hình của hiện tượng kinh tế - xã hội cũng bao gồm nhiều nhóm đơn vị, nhiều bộ phận có tính chất khác nhau. Cho nên muốn phản ánh được bản chất và quy luật phát triển của hiện tượng, phải nêu lên được đặc trưng của từng loại hình của từng bộ phận cấu thành hiện tượng phức tạp, đánh giá tầm quan trọng của mỗi bộ phận, nêu lên mối liên hệ giữa các bộ phận, từ đó mới nhận thức được đặc trưng của toàn bộ tổng thể hiện tượng nghiên cứu. Phân tổ thống kê giải quyết các nhiệm vụ cụ thể sau:

- Phân chia các loại hình kinh tế - xã hội của hiện tượng nghiên cứu:

Các loại hình của hiện tượng kinh tế - xã hội tồn tại khách quan. Sự vận động và phát triển của toàn bộ hiện tượng là kết quả đấu tranh giữa các loại hình đối lập tồn tại ngay trong bản thân hiện tượng. Cho nên phương pháp nghiên cứu khoa học là phải nêu rõ hiện tượng nghiên cứu bao gồm những loại hình nào. Muốn vậy phân tổ phải dựa vào lý luận kinh tế, chính trị, xã hội để phân biệt các bộ phận khác nhau về tính chất và tồn tại khách quan trong nội bộ hiện tượng. Tùy theo mục đích nghiên cứu và theo các tiêu thức khác nhau có thể phân chia hiện tượng kinh tế - xã hội thành các loại hình khác nhau. Loại phân tổ giải quyết nhiệm vụ cơ bản này gọi là *phân tổ phân loại*.

- Biểu hiện kết cấu của hiện tượng nghiên cứu:

Một loại hình kinh tế - xã hội, hoặc tổng thể đồng chất do nhiều bộ phận, nhiều nhóm đơn vị có tính chất khác nhau hợp thành. Các bộ phận hay nhóm này chiếm những tỷ trọng và biểu hiện tầm quan trọng khác nhau của chúng trong tổng thể. Tỷ trọng của các bộ phận còn biểu hiện kết cấu của tổng thể theo một tiêu thức nào đó. Muốn biểu hiện được kết cấu của tổng thể phân tổ thống kê phải xác định chính xác các bộ phận có tính chất khác nhau trong tổng thể, sau đó tính toán các tỷ trọng của từng bộ phận trong tổng thể. Loại phân tổ giải quyết nhiệm vụ này gọi là *phân tổ kết cấu*.

- Biểu hiện mối liên hệ giữa các tiêu thức:

Hiện tượng kinh tế - xã hội phát sinh và biến động một cách không ngẫu nhiên và không tách rời với các hiện tượng xung quanh, mà chúng có mối liên hệ và phụ thuộc lẫn nhau theo những quy luật nhất định. Giữa các tiêu thức mà thống kê nghiên cứu cũng thường có mối liên hệ với nhau. Sự thay đổi của tiêu thức này sẽ đưa đến sự thay đổi của tiêu thức kia theo một quy luật nhất định. Tìm hiểu tính chất và mức độ liên hệ giữa các hiện tượng nói chung và giữa các tiêu thức nói riêng là một trong các nhiệm vụ quan trọng của nghiên cứu thống kê.

Khi tiến hành phân tổ, các tiêu thức có liên quan với nhau được chia thành hai loại: Tiêu thức nguyên nhân (là tiêu thức gây nên ảnh hưởng) và tiêu thức kết quả (là tiêu thức chịu ảnh hưởng và phụ thuộc vào tiêu thức nguyên nhân). Loại phân tổ giải quyết nhiệm vụ này gọi là *phân tổ liên hệ*.

3.2.2. Các bước tiến hành phân tổ thống kê

3.2.2.1. Lựa chọn tiêu thức phân tổ

Tiêu thức phân tổ là tiêu thức được chọn làm căn cứ để tiến hành phân tổ thống kê. Lựa chọn tiêu thức phân tổ là vấn đề quan trọng phải được giải quyết chính xác. Các đơn vị tổng thể có nhiều tiêu thức khác nhau, tiêu thức nào cũng có thể dùng để phân tổ được. Song mỗi tiêu thức có ý nghĩa khác nhau, nêu lên đặc điểm khác nhau của hiện tượng. Có tiêu thức khi phân tổ phản ánh được bản chất của hiện tượng, có tiêu

thức chỉ nêu lên được một khía cạnh nào đó. Thậm chí có tiêu thức phân tổ không những không phản ánh được bản chất của hiện tượng mà còn làm cho người ta nhận thức sai lệch về hiện tượng.

Việc phân tổ chính xác và khoa học trước hết phụ thuộc vào việc lựa chọn tiêu thức phân tổ. Lựa chọn tiêu thức phân tổ chính xác, phản ánh đúng bản chất của hiện tượng nghiên cứu phải căn cứ vào các nguyên tắc sau đây:

- Phải dựa trên cơ sở phân tích lý luận một cách sâu sắc để chọn ra tiêu thức bản chất, phù hợp với mục đích nghiên cứu.

Tiêu thức bản chất là tiêu thức nêu lên được bản chất của hiện tượng nghiên cứu, phản ánh đặc trưng cơ bản của hiện tượng trong điều kiện thời gian và địa điểm cụ thể. Bản chất của hiện tượng có thể được phản ánh qua nhiều tiêu thức khác nhau, cho nên phải tùy mục đích nghiên cứu mà dùng lý luận để chọn ra tiêu thức bản chất nhất. Ví dụ: Khi phân tổ các doanh nghiệp thuộc ngành sản xuất công nghiệp nào đó, để nghiên cứu quy mô doanh nghiệp thì các tiêu thức bản chất là: Số lượng lao động, số vốn, giá trị sản xuất... Nhưng nếu nghiên cứu đơn vị tiên tiến hay lạc hậu thì các tiêu thức bản chất là: Giá thành sản phẩm, năng suất lao động, lợi nhuận...

- Phải căn cứ vào điều kiện lịch sử cụ thể của hiện tượng nghiên cứu để chọn ra tiêu thức phân tổ thích hợp.

Cùng loại hiện tượng nghiên cứu, nhưng phát sinh trong những điều kiện thời gian và địa điểm khác nhau, thì biểu hiện bản chất có thể khác nhau. Vì vậy, tiêu thức phân tổ cũng mang ý nghĩa khác nhau. Nếu chỉ dùng một tiêu thức chung cho mọi trường hợp, thì tiêu thức đó trong điều kiện này có thể giúp ta nghiên cứu chính xác, nhưng trong điều kiện khác thì không có tác dụng. Ví dụ khi nghiên cứu tình hình đời sống nông dân nước ta, trước kia có thể phân tổ các hộ gia đình theo thành phần giai cấp, theo số ruộng đất chiếm hữu... Nhưng hiện nay các tiêu thức phân tổ đó không phù hợp, nên phải chọn các tiêu thức phân tổ như: Số lao động, diện tích gieo trồng... là những tiêu thức phản ánh khả năng thu nhập và ảnh hưởng trực tiếp tới đời sống của nông dân.

- Tùy theo mục đích nghiên cứu và điều kiện tài liệu thực tế mà quyết định phân tổ hiện tượng theo một hay nhiều tiêu thức.

Phân tổ theo một tiêu thức gọi là phân tổ giản đơn, còn phân tổ theo nhiều tiêu thức gọi là phân tổ kết hợp. Nhìn chung, hiện tượng kinh tế - xã hội được nghiên cứu thường rất phức tạp, cho nên việc phân tổ theo một tiêu thức, dù là tiêu thức bản chất cũng chỉ phản ánh được một mặt nào đó của hiện tượng. Nếu phân tổ kết hợp theo nhiều tiêu thức sẽ phản ánh được nhiều mặt khác nhau của hiện tượng, các mặt này có thể bổ sung cho nhau và giúp cho việc nghiên cứu thêm sâu sắc. Trong nhiều trường hợp phân tổ kết hợp giúp ta nghiên cứu mối liên hệ giữa các tiêu thức. Ví dụ: Có thể phân tổ nhân khẩu kết hợp theo hai tiêu thức là tuổi và giới tính, phân tổ các doanh nghiệp kết hợp các tiêu thức theo nhóm, theo ngành và theo thành phần kinh tế... Tuy nhiên cũng không nên phân tổ kết hợp quá nhiều tiêu thức vì như vậy số tổ, tiểu tổ sẽ nhiều, tổng thể bị chia quá nhỏ, việc phân tổ phức tạp gây khó khăn cho việc nghiên cứu. Thường chỉ phân tổ kết hợp hai hay ba tiêu thức và nếu thấy cần thiết có thể lập các bảng phân tổ kết hợp khác nhau.

3.2.2.2. Xác định số tổ và khoảng cách tổ:

Sau khi chọn tiêu thức phân tổ thích hợp, vấn đề tiếp theo là xét xem cần phải chia hiện tượng nghiên cứu thành bao nhiêu tổ và căn cứ vào đâu để xác định số tổ cần thiết đó. Số tổ cần thiết thường được xác định tùy theo tiêu thức phân tổ là tiêu thức thuộc tính hay tiêu thức số lượng. Đối với mỗi loại tiêu thức thì việc xác định số tổ được giải quyết khác nhau.

** Phân tổ theo tiêu thức thuộc tính.*

Khi phân tổ theo tiêu thức thuộc tính, các tổ được hình thành không phải do sự khác nhau về số lượng biến của tiêu thức mà thường là do các loại hình khác nhau. Một số trường hợp phân tổ rất dễ dàng, vì số loại hình tương đối ít, như phân tổ nhân khẩu theo giới tính, phân tổ các doanh nghiệp theo thành phần kinh tế... ta có thể coi mỗi loại hình là một tổ.

Một số trường hợp phân tổ khác phức tạp như: Phân tổ nhân khẩu theo nghề nghiệp, phân tổ các mặt hàng theo giá trị sử dụng cụ thể, phân

tổ các ngành kinh tế quốc dân... Ở đây số loại hình thực tế rất nhiều có khi tới hàng trăm, hàng nghìn, nếu cứ coi mỗi loại hình là một tổ, tổng thể nghiên cứu bị chia nhỏ không giúp ta nghiên cứu được đặc trưng của tổng thể từ sự khác nhau của các tổ. Người ta phải ghép một số loại hình nhỏ vào một tổ theo nguyên tắc: Các loại hình nhỏ được ghép với nhau phải giống nhau (hoặc gần giống nhau) về tính chất nào đó hoặc về công dụng kinh tế - xã hội. Trên thực tế, người ta thường tiến hành sắp xếp và trình bày trong những văn bản gọi là bảng phân loại hay bảng danh mục do Nhà nước quy định thống nhất và cố định trong thời gian tương đối dài, nhằm đảm bảo tính chất so sánh được của tài liệu thống kê.

** Phân tổ theo tiêu thức số lượng.*

Khi phân tổ theo tiêu thức số lượng, phải căn cứ vào các lượng biến khác nhau của tiêu thức mà xác định các tổ khác nhau về tính chất. Tùy theo lượng biến của tiêu thức thay đổi nhiều hay ít mà phân tổ được giải quyết khác nhau. Mặt khác, cũng cần chú ý đến số lượng đơn vị tổng thể nhiều hay ít mà xác định số tổ thích hợp.

- Trường hợp lượng biến của tiêu thức thay đổi ít - nghĩa là sự biến thiên về mặt lượng giữa các đơn vị không chênh lệch nhau nhiều, biến động rời rạc và số các lượng biến ít như: Số người trong gia đình, số máy do công nhân phụ trách... thì số tổ có một giới hạn nhất định và thường cứ một lượng biến là cơ sở để hình thành một tổ. Ví dụ: Phân tổ các hộ gia đình của một địa phương theo số lượng nhân khẩu ta có kết quả ở bảng sau:

Bảng 3.1: Phân tổ số hộ gia đình theo số nhân khẩu

Số nhân khẩu	Số hộ
1	18
2	80
3	150
4	240
5	90
6	10

Việc phân tổ trên đây rất đơn giản, vì lượng biến của tiêu thức số nhân khẩu trong một hộ gia đình chỉ biến thiên trong phạm vi từ 1 người đến 6 người. Khi đó, mỗi một lượng biến hình thành một tổ. Phân tổ như trên gọi là phân tổ không có khoảng cách tổ.

- Trường hợp lượng biến của tiêu thức biến thiên lớn. Ta không thể áp dụng cách phân tổ trên được, nếu mỗi lượng biến hình thành một tổ thì số tổ sẽ quá nhiều và không nói rõ sự khác nhau về chất giữa các tổ. Trong trường hợp này ta cần chú ý mối liên hệ giữa lượng và chất trong phân tổ, xem lượng biến tích lũy đến một mức độ nào đó thì chất của lượng biến mới thay đổi và làm nảy sinh ra một số tổ khác. Như vậy, mỗi tổ sẽ bao gồm một phạm vi lượng biến, có hai giới hạn rõ rệt: *Giới hạn dưới* là lượng biến nhỏ nhất để làm cho tổ đó được hình thành và *giới hạn trên* là lượng biến lớn nhất của tổ, nếu vượt quá giới hạn này thì chất thay đổi mà chuyển sang tổ khác. Trị số chênh lệch giữa giới hạn trên và giới hạn dưới của mỗi tổ gọi là *khoảng cách tổ*. Việc phân tổ theo giới hạn như vậy gọi là phân tổ có khoảng cách tổ.

Để xác định ranh giới giữa các tổ, có hai trường hợp:

- Thứ nhất: Khi tiêu thức phân tổ biến thiên rời rạc thì giới hạn dưới của tổ nào đó là trị số sát với giới hạn trên của tổ đứng trước liền kề và giới hạn trên của tổ đó là trị số sát với giới hạn dưới của tổ đứng sau. Ví dụ: Phân tổ dân số theo nhóm tuổi, phân tổ các doanh nghiệp của một ngành sản xuất theo số lượng công nhân,...

- Thứ hai: Khi tiêu thức phân tổ biến thiên liên tục thì giới hạn dưới của một tổ nào đó là trị số trùng với giới hạn trên của tổ đứng trước liền kề và giới hạn trên của tổ đó là trị số trùng với giới hạn dưới của tổ đứng sau liền kề. Theo quy ước nếu một đơn vị nào đó có trị số tiêu thức trùng với giới hạn trên của tổ thì được xếp vào tổ kế tiếp. Ví dụ: Phân tổ các doanh nghiệp thương mại theo tỷ lệ phần trăm hoàn thành kế hoạch doanh thu, phân tổ số lao động trong một doanh nghiệp theo mức tiền lương,...

Ví dụ: Căn cứ theo tiêu thức số lượng công nhân để phân tổ các doanh nghiệp ngành sản xuất ở địa phương X như sau:

Bảng 3.2: Phân tổ các doanh nghiệp theo số lượng công nhân

Số công nhân	Số doanh nghiệp
< 100	25
101 - 200	70
201 - 500	55
501 - 1000	30
1001 - 2000	12
>2000	8

Trong phân tổ thống kê, các khoảng cách tổ có thể đều nhau hoặc không đều nhau. Việc xác định khoảng cách tổ đều nhau hay không đều nhau là phải căn cứ vào đặc điểm của hiện tượng nghiên cứu. Phân tổ phải đảm bảo các đơn vị được phân phối vào mỗi tổ đều có cùng một tính chất và sự khác nhau về lượng giữa các tổ phải nêu rõ được sự khác nhau về chất giữa các tổ đó. Trong thực tế, sự thay đổi về lượng của các bộ phận trong hiện tượng kinh tế - xã hội thường không diễn biến một cách đều đặn, bởi vì sự khác nhau về chất của chúng cũng không đều nhau. Do vậy nhiều trường hợp nghiên cứu phải phân tổ có khoảng cách tổ không đều nhau.

Đối với các hiện tượng tương đối đồng nhất về loại hình kinh tế - xã hội và lượng biến trên các đơn vị thay đổi tương đối đều đặn, có thể áp dụng phân tổ với khoảng cách đều nhau. Cách phân tổ này tạo điều kiện dễ dàng cho việc vận dụng các công thức toán học và trình bày số liệu trên các đồ thị thống kê.

Việc phân tổ với khoảng cách tổ đều nhau tương đối đơn giản và trị số khoảng cách tổ được xác định theo công thức:

$$h = \frac{X_{\max} - X_{\min}}{n}$$

Trong đó: h : Trị số khoảng cách tổ.

$X_{\max}$: Lượng biến lớn nhất của tiêu thức phân tổ.

$X_{\min}$: Lượng biến nhỏ nhất của tiêu thức phân tổ.

n : Số tổ dự định chia.

Trong một số trường hợp khi phân tổ mà tổ đầu tiên có thể không có giới hạn dưới, tổ cuối cùng không có giới hạn trên được gọi là *phân tổ mở*. Việc thành lập các tổ mở là cần thiết vì có tác dụng thu nạp đầy đủ các đơn vị có lượng biến nhỏ hoặc lớn bất thường, tránh việc hình thành quá nhiều tổ. Khi tính toán đối với tài liệu phân tổ mở, người ta quy ước lấy khoảng cách của tổ mở trên cơ sở khoảng cách tổ liền kề với nó.

3.2.2.3. *Các chỉ tiêu giải thích.*

Trong phân tổ thống kê, sau khi đã lựa chọn được tiêu thức phân tổ, xác định số tổ cần thiết còn phải xác định các chỉ tiêu giải thích để nói rõ đặc trưng của các tổ cũng như toàn bộ tổng thể. Chẳng hạn, sau khi phân tổ các doanh nghiệp theo thành phần kinh tế có thể đề ra một số chỉ tiêu giải thích theo ví dụ sau:

Bảng 3.3: Phân tổ các doanh nghiệp theo thành phần kinh tế

Thành phần kinh tế	Số DN	Số công nhân (người)	Giá trị tài sản cố định (triệu đồng)	Giá trị sản xuất (triệu đồng)	Năng suất lao động trung bình (tr.đ/ng)
- DN Nhà nước					
- DN Tập thể					
- DN Tư nhân					
- DN Liên doanh					
Chung					

Các chỉ tiêu: Số công nhân, giá trị tài sản cố định, giá trị sản xuất,... là các chỉ tiêu giải thích.

Mỗi chỉ tiêu giải thích có ý nghĩa riêng giúp ta thấy rõ các đặc trưng số lượng của từng tổ cũng như của toàn bộ tổng thể, làm căn cứ để so sánh các tổ với nhau và để tính ra một số chỉ tiêu phân tích khác. Tuy nhiên, cũng không nên đề ra quá nhiều chỉ tiêu, mà phải lựa chọn một số chỉ tiêu nào thích hợp nhất đối với mục đích nghiên cứu.

Muốn xác định các chỉ tiêu giải thích, chủ yếu phải căn cứ vào mục đích nghiên cứu và nhiệm vụ của phân tổ để chọn ra các chỉ tiêu có liên hệ với nhau và bổ sung cho nhau.

Các chỉ tiêu giải thích có ý nghĩa trong việc so sánh với nhau cần được bố trí gần nhau. Chẳng hạn nên bố trí chỉ tiêu thực hiện gần chỉ tiêu kế hoạch, chỉ tiêu tương đối gần chỉ tiêu tuyệt đối v.v...

3.2.3. Dãy số phân phối

Sau khi phân tổ tổng thể theo một tiêu thức nào đó, các đơn vị tổng thể được phân phối vào trong các tổ và ta sẽ có một dãy số phân phối. Đó là, dãy số trình bày có thứ tự số lượng đơn vị của từng tổ, trong một tổng thể đã được phân tổ theo một tiêu thức nhất định.

Dãy số phân phối có nhiều tác dụng trong nghiên cứu thống kê. Người ta thường dùng dãy số phân phối để khảo sát tình hình phân phối các đơn vị tổng thể theo một tiêu thức nghiên cứu, qua đó thấy được kết cấu của tổng thể và sự biến động kết cấu đó. Dãy số phân phối còn được dùng để tính ra nhiều chỉ tiêu nêu lên các đặc trưng của từng tổ và của tổng thể, biểu hiện mối liên hệ giữa các bộ phận hoặc giữa các tiêu thức. Tùy thuộc vào tiêu thức phân tổ người ta chia dãy số phân phối thành hai loại:

Dãy số phân phối theo tiêu thức thuộc tính (còn gọi là dãy số thuộc tính): Dãy số này phản ánh kết cấu của tổng thể theo một tiêu thức thuộc tính nào đó. Ví dụ dãy số phân phối các doanh nghiệp sản xuất công nghiệp theo thành phần kinh tế, theo ngành sản xuất v.v...

Dãy số phân phối theo tiêu thức số lượng (còn gọi là dãy số lượng biến): Dãy số này phản ánh kết cấu của tổng thể theo một tiêu thức số lượng nào đó. Ví dụ: Dãy số phân phối một tổng thể công nhân theo mức lương, dãy số phân phối nhân khẩu theo độ tuổi v.v...

Một dãy số lượng biến có hai thành phần: Lượng biến và tần số. Lượng biến là các trị số nói lên biểu hiện cụ thể của tiêu thức số lượng, thường ký hiệu bằng x_i . Tần số là số đơn vị được phân phối vào trong mỗi tổ, tức là số lần một lượng biến nhận một trị số nhất định trong tổng thể. Tần số thường được ký hiệu bằng f_i . Khi tần số được biểu hiện bằng

số tương đối (số %), gọi là tần suất thường được ký hiệu là d_i . Tần số hoặc tần suất được sắp xếp theo trình tự của lượng biến thường là từ nhỏ đến lớn. Khi cần quan sát đơn vị thứ bao nhiêu từ trên xuống ứng với lượng biến nào, người ta cộng dồn tần số từ trên xuống, các tổng đó gọi là tần số tích lũy tiến, ngược lại cộng dồn từ dưới lên được gọi là tần số tích lũy lùi.

Dạng chung của một dãy số lượng biến như sau:

Lượng biến (x_i)	Tần số (f_i)	Tần số tích lũy tiến	Tần số tích lũy lùi
x_1	f_1	f_1	$f_n + f_n + \dots + f_1$
x_2	f_2	$f_1 + f_2$	$\dots$
x_3	f_3	$f_1 + f_2 + f_3$	$f_n + f_n + \dots + f_3$
.	.	.	
.	.	.	
.	.	.	
x_{n-1}	f_{n-1}	$f_1 + \dots f_{n-1}$	$f_n + f_n - 1$
x_n	f_n	$f_1 + \dots f_n$	f_n

Lượng biến thiên của tiêu thức số lượng chia làm hai loại: Lượng biến không liên tục và lượng biến liên tục. Lượng biến không liên tục (còn gọi là lượng biến rời rạc) chỉ có các trị số bằng số nguyên như số công nhân, số hợp tác xã, số xe vận tải... Lượng biến liên tục có thể được biểu hiện bằng những trị số bất kỳ (số nguyên hoặc số thập phân) như năng suất lúa, tỷ lệ hoàn thành kế hoạch...

Nếu dãy số phân phối theo tiêu thức số lượng có các tổ với khoảng cách tổ không bằng nhau thì tần số trong các khoảng cách tổ không trực tiếp so sánh với nhau được, vì các tần số đó phụ thuộc vào trị số khoảng cách tổ. Để có thể so sánh được các tần số, người ta tính mật độ phân phối. Mật độ phân phối của mỗi tổ là tỷ số giữa tần số (hoặc tần suất) với trị số khoảng cách tổ của tổ đó. Ta có công thức sau:

$$m_i = \frac{f_i}{h_i}$$

Trong đó: m_i : Mật độ phân phối

f_i : Tần số

h_i : Khoảng cách tổ

3.3. Bảng thống kê và đồ thị thống kê

3.3.1. Bảng thống kê

3.3.1.1. Ý nghĩa tác dụng của bảng thống kê

Bảng thống kê là một hình thức biểu hiện các tài liệu thống kê một cách có hệ thống, hợp lý và rõ ràng, nhằm nêu lên các đặc trưng về mặt lượng của hiện tượng nghiên cứu. Đặc điểm chung của tất cả các bảng thống kê bao giờ cũng có những con số bộ phận là chung, các con số này có liên hệ mật thiết với nhau.

Bảng thống kê có nhiều tác dụng quan trọng trong mọi công tác nghiên cứu kinh tế - xã hội. Các tài liệu trong bảng thống kê đã được sắp xếp một cách khoa học, nên có thể giúp ta tiến hành việc so sánh, đối chiếu, phân tích theo các phương pháp khác nhau, nhằm nêu lên một cách sâu sắc bản chất của hiện tượng nghiên cứu. Nếu biết trình bày và sử dụng thích đáng các bảng thống kê thì việc chứng minh vấn đề sẽ trở nên rất sinh động, có sức thuyết phục hơn cả những bài văn dài.

3.3.1.2. Cấu thành bảng thống kê

- *Về hình thức*: Bảng thống kê bao gồm các hàng ngang và cột dọc, các tiêu đề và các tài liệu con số.

Các hàng ngang, cột dọc phản ánh quy mô của bảng thống kê vì số hàng và cột càng nhiều thì bảng thống kê càng lớn và phức tạp. Các hàng ngang, cột dọc cắt nhau tạo thành các ô dùng để điền các số liệu thống kê vào đó. Các hàng và cột thường được đánh số thứ tự để tiện cho việc sử dụng và trình bày vấn đề.

Tiêu đề của bảng phản ánh nội dung của bảng và của từng chi tiết trong bảng. Trước hết có tiêu đề chung, là tên gọi chung của bảng, thường viết ngắn, gọn, dễ hiểu và đặt ở phía trên đầu của bảng. Các tiêu

đề nhỏ (còn gọi là tiêu mục) là tên riêng của mỗi hàng và cột phản ánh rõ nội dung các hàng và cột đó.

Các tài liệu con số được ghi vào các ô của bảng, mỗi con số phản ánh một đặc trưng về mặt lượng của hiện tượng nghiên cứu.

- Về nội dung: Bảng thống kê gồm hai phần: Phần chủ đề và phần giải thích.

Phần chủ đề (hay còn gọi là phần chủ từ) nêu lên tổng thể hiện tượng được trình bày trong bảng thống kê, tổng thể này được phân chia thành những bộ phận nào. Nó giải đáp đối tượng nghiên cứu là những đơn vị nào, những loại hình gì. Có khi phần chủ đề là các địa phương hoặc thời gian nghiên cứu khác nhau của hiện tượng nào đó.

Phần giải thích (còn gọi là phần tân từ) gồm các chỉ tiêu giải thích các đặc điểm của đối tượng nghiên cứu, nghĩa là giải thích phần chủ đề của bảng.

Phần chủ đề thường được đặt ở vị trí bên trái của bảng, còn phần giải thích ở phía trên của bảng. Cũng có trường hợp người ta thay đổi vị trí của hai phần cho nhau.

Cấu thành của bảng thống kê có thể biểu hiện khái quát như sau:

Tên bảng thống kê (tiêu đề chung)				
Phần giải thích Phần chủ đề	Các chỉ tiêu giải thích (tên cột)			
A	(1)	(2)	(...)	(n)
Tên chủ đề (Tên hàng)				
Tổng cộng	↑	↑	↑	↑

Số liệu
các cột

Các
hàng
của
bảng

Hàng chung

Các cột của bảng

Cột chung

3.3.1.3. Các loại bảng thống kê

Căn cứ theo kết cấu của phần chủ đề, có thể chia làm ba loại bảng thống kê: Bảng giản đơn, bảng phân tổ và bảng kết hợp.

a. Bảng giản đơn

Bảng giản đơn là loại bảng thống kê, trong đó phần chủ đề không phân tổ. Trong phần chủ đề của bảng giản đơn có liệt kê các đơn vị tổng thể, tên gọi các địa phương hoặc các thời gian khác nhau của quá trình nghiên cứu. Ví dụ có bảng giản đơn sau:

**Bảng 3.4: Tình hình sản xuất kinh doanh năm N
của các doanh nghiệp thuộc một ngành X**

Tên doanh nghiệp	Số lao động	Giá trị sản xuất (1000đ)	Năng suất lao động bình quân
1	2	3	4
Doanh nghiệp A			
Doanh nghiệp B			
Doanh nghiệp C			
...			
Cộng			

b. Bảng phân tổ

Bảng phân tổ là loại bảng thống kê, trong đó đối tượng nghiên cứu ghi trong phần chủ đề được phân chia thành các tổ theo một tiêu thức nào đó. Các bảng phân tổ là kết quả của việc áp dụng phương pháp phân tổ thống kê. Bảng phân tổ cho ta thấy rõ các loại hình kinh tế - xã hội tồn tại trong bản thân hiện tượng nghiên cứu, nêu lên kết cấu và biến động kết cấu của hiện tượng, trong nhiều trường hợp còn giúp ta phân tích được mối liên hệ giữa các hiện tượng.

c. Bảng kết hợp

Bảng kết hợp là loại bảng thống kê, trong đó đối tượng nghiên cứu ghi trong phần chủ đề được phân tổ theo hai, ba,... tiêu thức kết hợp với

nhau. Loại bảng kết hợp giúp ta nghiên cứu được sâu sắc bản chất của hiện tượng, đi sâu vào kết cấu nội bộ của hiện tượng, thấy rõ mối quan hệ giữa các tổ, bộ phận của hiện tượng trong quá trình phát triển.

3.3.1.4. Các yêu cầu đối với việc xây dựng bảng thống kê

Một bảng thống kê được xây dựng một cách khoa học sẽ trở nên gọn, rõ, đáp ứng được mục đích nghiên cứu. Việc xây dựng bảng thống kê cần đảm bảo những yêu cầu sau:

- *Thứ nhất*, quy mô của bảng thống kê không nên quá lớn, tức là quá nhiều hàng, cột và nhiều phân tổ kết hợp. Một bảng thống kê ngắn, gọn một cách hợp lý sẽ tạo điều kiện dễ dàng cho việc phân tích. Nếu thấy cần thiết nên xây dựng hai, ba,... bảng thống kê nhỏ thay cho một bảng quá lớn.

- *Thứ hai*, các tiêu đề và tiêu mục trong bảng thống kê cần được ghi chính xác, gọn và dễ hiểu. Tiêu đề chung không những nói rõ nội dung chủ yếu của bảng thống kê, mà còn cần chỉ rõ hiện tượng nghiên cứu vào thời gian và địa điểm nào? Nhiều khi ở phần tiêu đề chung còn quy định đơn vị tính toán chung cho các số liệu trong bảng thống kê (nếu đơn vị tính toán không thống nhất cho các số liệu, thì chỉ quy định riêng cho mỗi hàng và cột).

- *Thứ ba*, các hàng và cột thường được ký hiệu bằng chữ hoặc bằng số để tiện cho việc trình bày hoặc giải thích nội dung. Các cột của phần chủ đề thường được ký hiệu bằng các chữ a, b, c,... còn các cột của phần giải thích được ký hiệu bằng số 1, 2, 3,... Tuy nhiên, nếu một bảng thống kê chỉ có ít hàng, cột và nội dung các hàng, cột đã rõ ràng, dễ hiểu thì không nhất thiết phải dùng ký hiệu.

- *Thứ tư*, các chỉ tiêu giải thích trong bảng thống kê cần được sắp xếp theo thứ tự hợp lý, phù hợp với mục đích nghiên cứu. Giả sử muốn lập một bảng thống kê nêu rõ mối liên hệ giữa mức năng suất lao động và giá trị sản lượng của các xí nghiệp. Như vậy, trước hết trong phần chủ đề ta có thể phân tổ các xí nghiệp theo giá trị sản lượng từ thấp đến cao (phân tổ có khoảng cách tổ). Còn các chỉ tiêu giải thích được bố trí theo thứ tự sau: Số xí nghiệp mỗi tổ, giá trị sản lượng của các xí nghiệp trong

tổ, số công nhân bình quân trong kỳ của mỗi tổ, năng suất lao động bình quân của mỗi công nhân trong tổ. Nếu bây giờ ta đảo ngược trật tự các chỉ tiêu nói trên, thì việc nhận thức và tính toán sẽ khó khăn hơn.

Trong mỗi bảng thống kê, các chỉ tiêu có ý nghĩa quan trọng trong việc so sánh với nhau thì nên bố trí gần nhau, như chỉ tiêu thực hiện bố trí gần chỉ tiêu kế hoạch, chỉ tiêu tương đối bố trí gần chỉ tiêu tuyệt đối,...

- Thứ năm, cách ghi các số liệu vào bảng thống kê: Các ô trong bảng thống kê đều có ghi số liệu hoặc bằng các ký hiệu quy ước thay thế. Thường dùng các ký hiệu quy ước sau:

+ Nếu hiện tượng không có số liệu đó, thì trong ô sẽ ghi một dấu gạch ngang (-).

+ Nếu số liệu còn thiếu, sau này có thể bổ sung, thì trong ô có ký hiệu 3 chấm (...).

+ Ký hiệu gạch chéo (x) trong một ô nào đó nói lên rằng hiện tượng không có liên quan đến chỉ tiêu đó, nếu viết số liệu vào ô đó sẽ vô nghĩa.

Các số liệu trong cùng một cột, có đơn vị tính toán giống nhau, phải ghi theo trình độ chính xác như nhau (số lẻ đến 0,1 hay 0,01), đơn vị tính phải ghi thống nhất theo quy định.

Nếu mục đích của bảng thống kê chỉ nhằm nêu lên những nét chung về bản chất hiện tượng, không cần quá chi li số lẻ thì các số liệu trong bảng có thể ghi theo số tròn. Chẳng hạn, các đơn vị đo lường tính lẻ đến kilogram có thể tính tròn đến tạ, tấn; đơn vị đo lường tính lẻ đến từng mét có thể tính tròn đến kilômét; tiền tệ có thể tính tròn đến nghìn hoặc triệu đồng,... Bằng cách tính tròn như vậy, có thể thay những số liệu có 6, 7,... chữ số thành những số liệu chỉ có gọn 2, 3, ... chữ số. Việc tính tròn cũng theo nguyên tắc toán học.

Các số cộng và tổng cộng có thể được ghi ở đầu hoặc ở cuối hàng và cột tùy theo mục đích nghiên cứu. Các số này được ghi ở đầu bảng, đầu cột khi ta cần nghiên cứu chủ yếu các đặc trưng của hiện tượng, còn các đặc trưng của từng bộ phận chỉ có tác dụng phân tích thêm. Các số

cộng và tổng được ghi ở cuối hàng, cuối cột là khi ta nghiên cứu đi sâu từng tổ, từng bộ phận là chủ yếu.

- Thứ sáu, phần ghi chú ở cuối bảng thống kê được dùng để giải thích rõ nội dung của một số chỉ tiêu trong bảng, để nói rõ nguồn số liệu đã được sử dụng trong bảng hoặc các chi tiết cần thiết khác.

3.3.2. Đồ thị thống kê

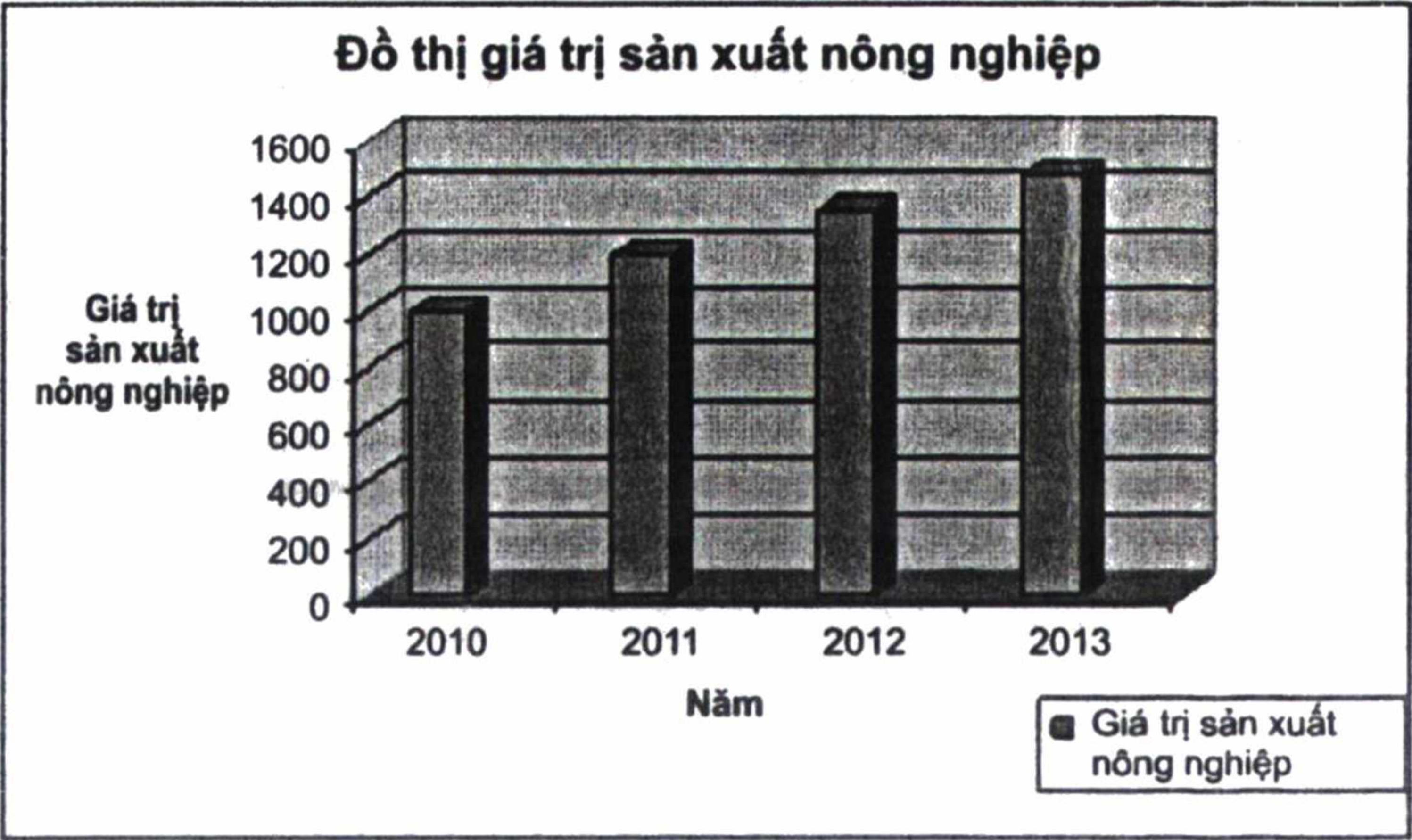
3.3.2.1. Ý nghĩa và tác dụng của đồ thị thống kê

Đồ thị thống kê là các hình vẽ hoặc đường nét hình học dùng để miêu tả có tính chất quy ước các tài liệu thống kê. Ví dụ có tài liệu về tốc độ phát triển giá trị sản xuất nông nghiệp của địa phương như sau:

Chỉ tiêu	Năm 2010	Năm 2011	Năm 2012	Năm 2013
Giá trị sản xuất nông nghiệp (tỷ đồng)	1000	1200	1350	1480

Với tài liệu trên, để dễ thấy, dễ hiểu sự phát triển của giá trị sản xuất nông nghiệp, ta có thể dùng hình vẽ sau:

Đồ thị biểu hiện giá trị sản xuất nông nghiệp



Trong hình vẽ trên, các cột đứng nói lên sự phát triển của sản xuất nông nghiệp của địa phương qua các năm. Với cùng một gốc và cùng một bề rộng, các chiều cao khác nhau của mỗi hình cột cho ta một nhận thức có so sánh cụ thể về sự phát triển của hiện tượng nghiên cứu. Các hình vẽ gọi là biểu đồ thống kê, còn phương pháp dùng các hình vẽ để miêu tả hiện tượng qua các con số thống kê gọi là phương pháp đồ thị thống kê.

Đồ thị thống kê có các đặc điểm chủ yếu sau:

- Đồ thị thống kê sử dụng con số kết hợp với hình vẽ, đường nét và màu sắc để trình bày và phân tích các đặc trưng số lượng của hiện tượng. Vì vậy, người xem không mất nhiều công đọc con số mà vẫn nhận thức được vấn đề chủ yếu một cách dễ dàng, nhanh chóng.

- Đồ thị thống kê chỉ trình bày một cách khái quát các đặc điểm chủ yếu về bản chất và xu hướng phát triển của các hiện tượng.

Do các đặc điểm nêu trên, đồ thị thống kê có tính quần chúng, có sức hấp dẫn và sinh động, làm cho người hiểu biết ít về thống kê vẫn lĩnh hội được vấn đề chủ yếu một cách dễ dàng, đồng thời giữ được ấn tượng khá sâu đối với hiện tượng.

Ngoài ra, đồ thị thống kê còn được coi là một phương tiện tuyên truyền, một công cụ dùng để biểu dương các kết quả sản xuất và hoạt động văn hóa - xã hội.

3.3.2.2. Phân loại đồ thị thống kê

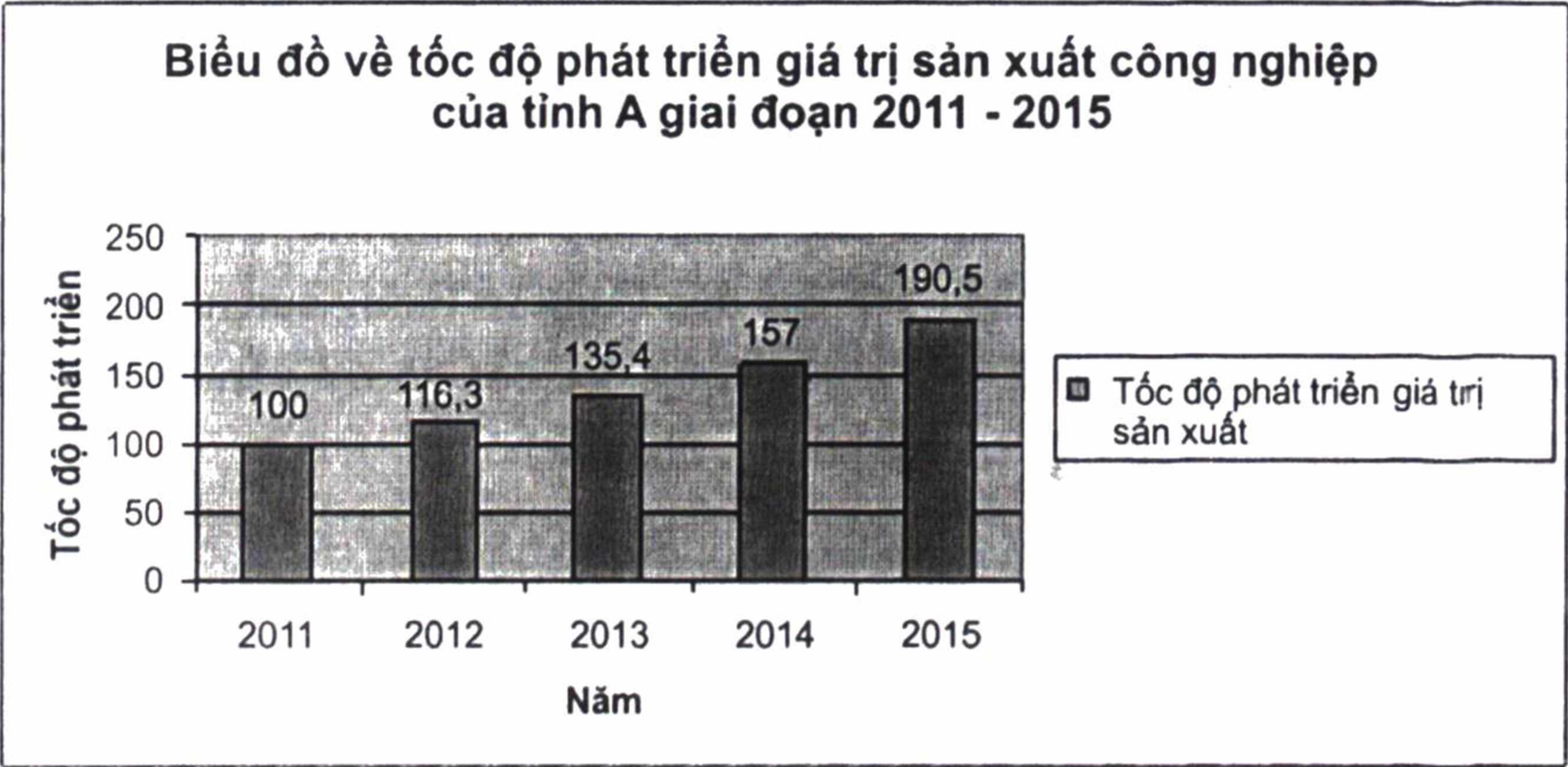
- * Căn cứ vào hình thức biểu hiện, có thể phân chia đồ thị thống kê thành các loại sau: Đồ thị hình cột, đồ thị hình tượng, đồ thị diện tích (tròn, chữ nhật, vuông), đồ thị đường gấp khúc, đồ thị ra đa.

- * Căn cứ vào nội dung phản ánh, có thể phân chia đồ thị thống kê thành các loại sau:

- Đồ thị phát triển: Đồ thị này dùng để biểu hiện tình hình phát triển của hiện tượng và so sánh giữa các hiện tượng, có thể dùng các loại đồ hình cột, hình tròn và đồ thị tuyến tính.

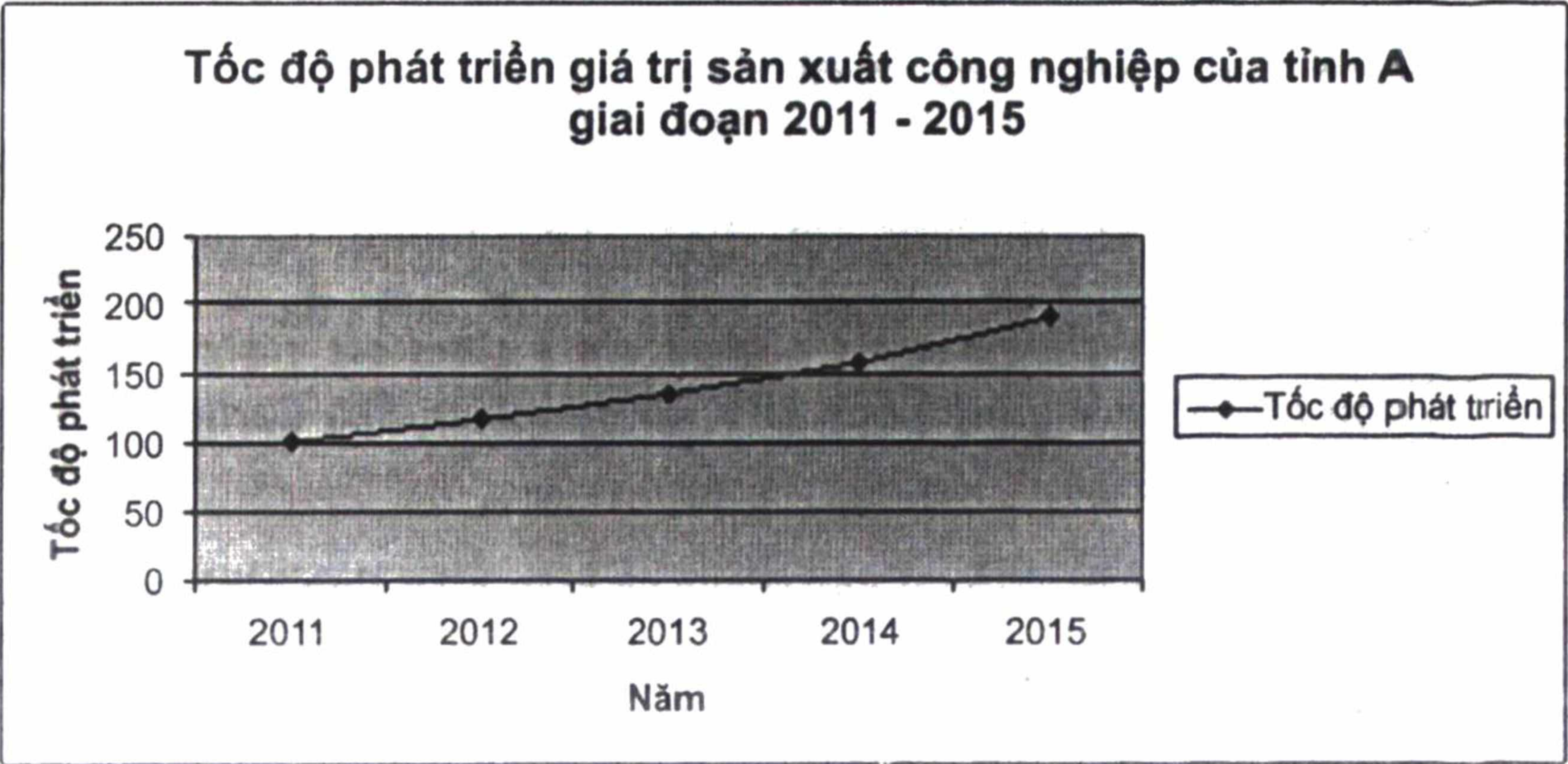
Ví dụ: Có tài liệu về tốc độ phát triển giá trị sản xuất công nghiệp của tỉnh A giai đoạn 2011 - 2015 như sau (lấy 2011 làm gốc):

Năm	2011	2012	2013	2014	2015
Tốc độ phát triển	100	116,3	135,4	157	190,5



Trong đồ thị trên, các cột đứng nói lên sự phát triển của sản xuất công nghiệp tỉnh A giai đoạn 2011 - 2015. Các cột có bề rộng bằng nhau, còn chiều cao tương ứng với các đại lượng được biểu hiện.

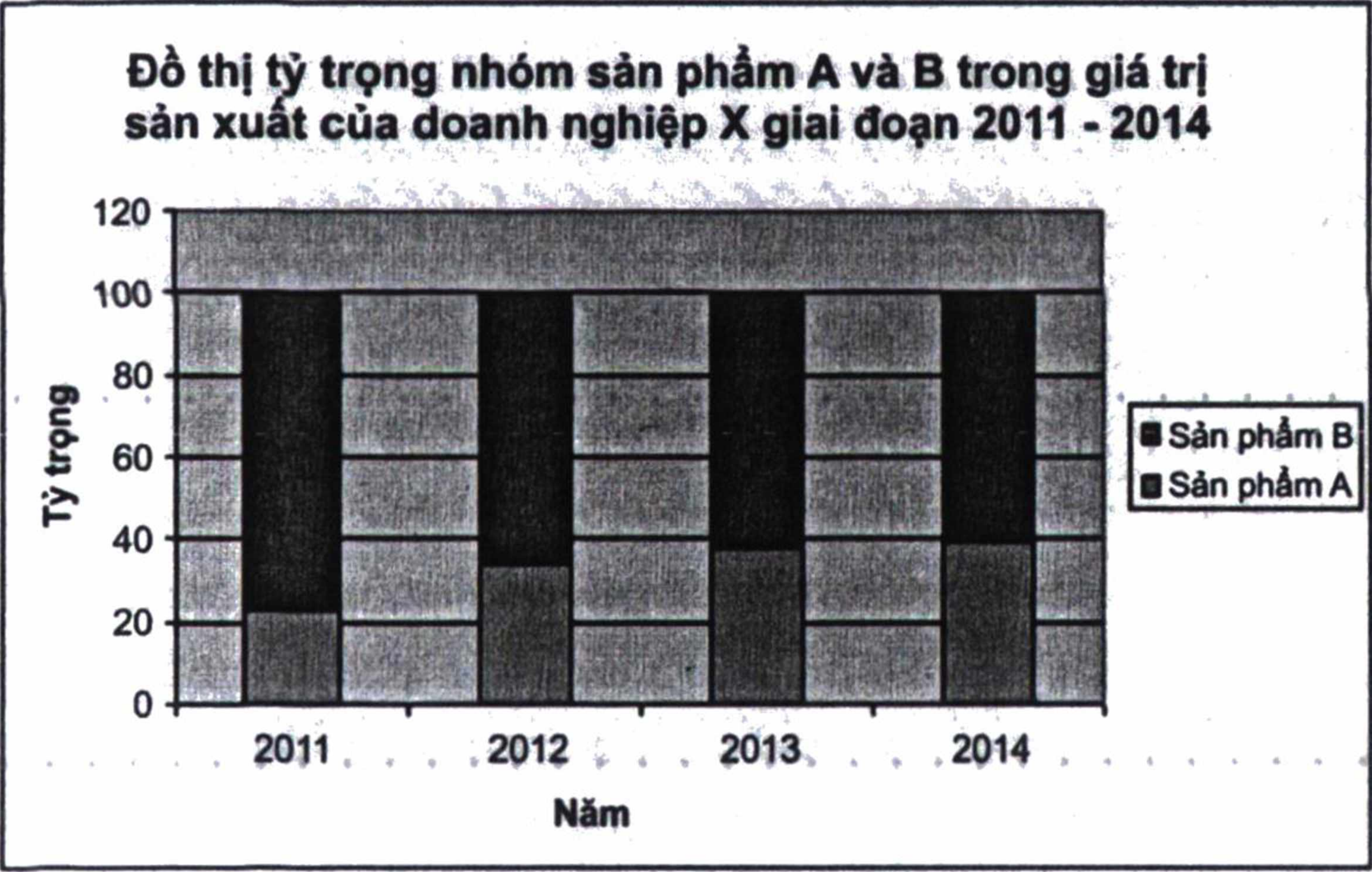
Các đồ thị tuyến tính cũng thường được dùng để biểu hiện sự phát triển của hiện tượng, theo ví dụ trên ta có thể vẽ thành đồ thị sau:



Trên đồ thị tuyến tính, trục hoành thường được dùng để biểu hiện thời gian, còn trục tung biểu hiện các mức độ của chỉ tiêu nghiên cứu. Một chú ý quan trọng khi vẽ các loại đồ thị này là phải xác định độ khắc trên các trục tọa độ cho thích hợp, vì độ khắc có ảnh hưởng trực tiếp đến độ dốc của đường gấp khúc. Nếu độ khắc trên trục tung quá nhỏ so với độ khắc trên trục hoành, đường gấp khúc sẽ vươn dài một cách quá mức, độ dốc của đường sẽ không thấy rõ. Ngược lại, nếu độ khắc trên trục tung quá lớn so với độ khắc trên trục hoành, đường gấp khúc sẽ vươn cao quá mức, độ dốc quá lớn gây cho người xem ấn tượng phóng đại về sự phát triển của hiện tượng.

- Đồ thị kết cấu: Đồ thị này dùng để biểu hiện kết cấu và biến động kết cấu của hiện tượng, thường dùng các loại đồ thị hình cột, hình tròn (có thể chia nhỏ thành các hình quạt)

Ví dụ: Có đồ thị hình cột về kết cấu giá trị sản xuất của một doanh nghiệp giai đoạn 2011 - 2014, với 2 loại sản phẩm A và B như sau:



- Đồ thị liên hệ: Dùng để biểu hiện mối liên hệ giữa 2 tiêu thức, người ta thường dùng đồ thị đường gấp khúc. Trục hoành của đồ thị thường biểu hiện trị số của tiêu thức nguyên nhân (tiêu thức gây ảnh hưởng), trục tung biểu hiện trị số của tiêu thức kết quả (tiêu thức chịu ảnh hưởng).

3.3.2.3. Những yêu cầu chung đối với việc xây dựng đồ thị thống kê

Một đồ thị thống kê phải đảm bảo các yêu cầu sau: Chính xác, dễ xem, dễ hiểu và có tính mỹ thuật, để đảm bảo những yêu cầu này ta phải chú ý đến các yếu tố chính của đồ thị, quy mô, các ký hiệu hình học hoặc các hình vẽ, hệ tọa độ, thang và tỷ lệ xích, phần giải thích.

Quy mô của đồ thị được quyết định bởi chiều dài, chiều cao và quan hệ tỷ lệ giữa hai chiều đó. Quy mô của đồ thị to hay nhỏ còn phải căn cứ vào mục đích sử dụng, trong các báo cáo phân tích không nên vẽ các đồ thị quá lớn. Quan hệ tỷ lệ giữa chiều cao và chiều dài của đồ thị thông thường được dùng từ 1:1,33 đến 1:1,15.

Các ký hiệu hình học hoặc hình vẽ quyết định hình dáng của đồ thị. Các ký hiệu hình học thường có nhiều loại như: Các chấm, các đường thẳng hoặc cong, các hình cột, hình vuông, hình chữ nhật, hình tròn... Các hình vẽ khác trên đồ thị cũng có thể thay đổi nhiều loại tùy thuộc vào tính chất của hiện tượng nghiên cứu. Việc lựa chọn các ký hiệu hình học hoặc hình vẽ của đồ thị là vấn đề quan trọng, vì mỗi hình có khả năng diễn tả riêng. Ví dụ, khi cần biểu hiện kết cấu của hiện tượng nghiên cứu, có thể vẽ các hình cột (có chia thành nhiều đoạn) hoặc các hình tròn (có chia thành các hình quạt), hoặc hình vuông, hình chữ nhật... Nhưng người ta thường dùng hình tròn, vì loại này biểu hiện được rõ nhất kết cấu và biến động kết cấu của hiện tượng.

Hệ tọa độ giúp cho việc xác định chính xác vị trí các ký hiệu hình học trên đồ thị. Các đồ thị thống kê thường dùng hệ tọa độ vuông góc. Trong các đồ thị thống kê, người ta thường dùng các đường cong để làm căn cứ xác định vị trí các ký hiệu hình học. Các đường cong này có thể là đường biên giới, đường dọc bờ biển, các con sóng lớn... trên hệ tọa độ vuông góc, trục hoành thường được dùng để biểu thị thời gian, trục tung biểu thị trị số của chỉ tiêu. Trong trường hợp phân tích mối liên hệ giữa hai tiêu thức, thì tiêu thức nguyên nhân được để ở trục hoành, tiêu thức kết quả được ghi trên trục tung.

Thang và tỷ lệ xích giúp cho việc tính chuyển các đại lượng lên đồ thị theo các khoảng cách thích hợp. Người ta thường dùng các thang

đường thẳng, được phân bố theo các trục tọa độ. Cũng có khi dùng thang đường cong, ví dụ thang tròn (ở đồ thị hình tròn) được chia thành 360° . Các thang đo tỷ lệ có thể có khoảng cách bằng nhau hoặc không bằng nhau. Các thang đo tỷ lệ có khoảng cách không bằng nhau như thang lôgarit chỉ dùng để biểu hiện các tốc độ khi khoảng biến thiên của các mức độ quá lớn mà người ta chỉ chú ý đến biến động tương đối của chúng.

Phân giải thích bao gồm tên đồ thị, các con số và ghi chú dọc theo thang tỷ lệ, các con số bên cạnh từng bộ phận của đồ thị, giải thích các ký hiệu quy ước... cần được ghi rõ, gọn, dễ hiểu.

TÓM TẮT CHƯƠNG 3

Tổng hợp thống kê là tiến hành tập trung chỉnh lý và hệ thống hóa một cách khoa học các tài liệu thu thập được trong điều tra thống kê. Đây là giai đoạn thứ hai của quá trình nghiên cứu thống kê với mục tiêu cơ bản là làm cho những đặc trưng riêng biệt trên từng đơn vị bước đầu chuyển thành đặc trưng của toàn bộ tổng thể nghiên cứu. Trong trường hợp khối lượng tài liệu tổng hợp gồm nhiều đơn vị để tiến hành tổng hợp, thống kê sử dụng phương pháp phân tổ thống kê. Quá trình phân tổ thống kê được thực hiện theo ba bước: Xác định tiêu thức phân tổ, xác định số tổ cần thiết và khoảng cách tổ, xác định tổ trên giải thích và sắp đến các đơn vị vào từng tổ.

Sau khi tiến hành phân tổ thống kê theo một hay một số tiêu thức nào đó, các đơn vị được phân phối vào các tổ, ta sẽ có một dãy số phân phối. Có hai loại dãy số phân phối: dãy số phân phối theo tiêu thức thuộc tính và dãy số phân phối theo tiêu thức số lượng.

Kết quả của quá trình phân tổ thống kê có thể được trình bày bằng bảng thống kê hoặc đồ thị thống kê. Bảng thống kê là hình thức trình bày số liệu một cách chi tiết, có hệ thống. Về hình thức một bảng thống kê bao gồm các hàng ngang, cột dọc, tiêu đề, tiêu mục và con số. Có ba loại bảng thống kê: Bảng giản đơn, bảng phân tổ và bảng kết hợp. Khi xây dựng bảng thống kê cần đảm bảo một số yêu cầu nhất định. Đồ thị thống

kê là hình thức mô tả tài liệu thống kê một cách quy ước bằng cách sử dụng các hình vẽ hoặc đường nét hình học. Đồ thị thống kê được sử dụng rộng rãi trong nghiên cứu kinh tế. Có nhiều loại đồ thị thống kê như biểu đồ hình tròn, biểu đồ diện tích, biểu đồ hình tượng. Mỗi loại biểu đồ có khả năng diễn đạt riêng, do vậy trong mỗi trường hợp cụ thể cần phải lựa chọn sử dụng loại biểu đồ thích hợp.

CÂU HỎI ÔN TẬP CHƯƠNG 3

- 1. Trình bày khái niệm, ý nghĩa và nhiệm vụ của tổng hợp thống kê, những vấn đề cơ bản của tổng hợp thống kê?
- 2. Trình bày khái niệm, ý nghĩa và nhiệm vụ của phân tổ thống kê?
- 3. Trình bày khái niệm tiêu thức phân tổ thống kê và phân tích các nguyên tắc lựa chọn tiêu thức phân tổ thống kê?
- 4. Trình bày các bước phân tổ thống kê? Cho ví dụ minh họa?
- 5. Trình bày đặc điểm, tác dụng của bảng thống kê, các loại bảng thống kê và các nguyên tắc trình bày bảng thống kê?
- 6. Trình bày đặc điểm tác dụng, phân loại đồ thị thống kê? Cho ví dụ minh họa về một số loại đồ thị thống kê?

BÀI TẬP CHƯƠNG 3

Bài 1:

Có tài liệu về bậc thợ của 40 công nhân trong một doanh nghiệp như sau:

Bậc thợ công nhân:

1 2 3 4 3 1 2 7 1 3
2 4 3 4 5 1 3 3 2 4
6 2 6 3 3 4 3 2 4 3
3 5 2 3 1 4 5 1 2 3

Hãy phân tổ tài liệu trên đây nhằm phản ánh tình hình phân phối số công nhân của doanh nghiệp theo bậc thợ. Trình bày kết quả bằng bảng thống kê. Cho nhận xét?

Bài 2:

Có tài liệu về năng suất lao động của các nhân viên bán hàng trong công ty X như sau:

Năng suất lao động (trđ)

28,7	31,0	33,0	35,0	29,0	33,9
33,7	29,4	31,8	36,9	28,0	34,3
29,8	32,1	33,8	35,2	34,9	36,6
30,2	32,4	34,8	28,9	33,1	38,9
37,2	36,7	30,4	32,9	34,6	43,0

Hãy phân tổ các nhân viên theo năng suất lao động thành 5 tổ có khoảng cách tổ. Trình bày kết quả phân tổ bằng bảng thống kê. Cho nhận xét?

Bài 3:

Có tài liệu phân tổ công nhân theo mức lương tại hai phân xưởng như sau:

Phân xưởng A		Phân xưởng B	
Mức lương (tr.đ)	Tỷ lệ công nhân so với tổng số (%)	Mức lương (tr.đ)	Tỷ lệ công nhân So với tổng số (%)
5-6	5	4-5,5	4
6-7	12	5,5-7	10
7-8	18	7-8,5	30
8-9	26	8,5-10	21
9-10	25	10-11,5	15
10-11	7	11,5-13	16
11-12	4	13-14,5	4
12-13	3		
Cộng	100	Cộng	100

Hãy phân tổ lại các tài liệu trên đây, nhằm mục đích so sánh tiền lương của công nhân hai phân xưởng (chia thành 4 tổ: Dưới 7, từ 7 đến 10, từ 10 đến 13, trên 13).

Bài 4:

Hãy phân tổ 2000 sinh viên ở một khoá học trong trường đại học, theo tiêu thức điểm bình quân tích lũy, thành 4 tổ có khoảng cách không đều. Trình bày kết quả bằng bảng thống kê. Cho nhận xét?

Bài 5:

Doanh nghiệp X có 1000 công nhân. Anh (chị) hãy phân tổ công nhân theo tiêu thức “thu nhập trung bình 1 công nhân trong tháng”, thành 3 tổ, có khoảng cách tổ mở. Lập bảng thống kê. Cho nhận xét?

Chương 4

THỐNG KÊ MỨC ĐỘ CỦA HIỆN TƯỢNG KINH TẾ - XÃ HỘI

Đặc điểm của hiện tượng kinh tế - xã hội được biểu hiện bằng các mức độ khác nhau. Các mức độ cho ta một nhận thức cụ thể về quy mô, khối lượng của hiện tượng nghiên cứu, phản ánh các quan hệ tỷ lệ khác nhau. Thông qua nghiên cứu các mức độ có thể thấy được đặc điểm chung nhất, đại diện nhất về từng mặt của hiện tượng nghiên cứu bao gồm nhiều đơn vị cùng loại, cũng như đánh giá mức độ đồng đều và phân phối các đơn vị tổng thể. Đây là những cơ sở dữ liệu quan trọng nhất để phân tích và dự báo thống kê. Nội dung chương này sẽ trình bày về lý luận và phương pháp tính các mức độ của hiện tượng bao gồm: Số tuyệt đối, số tương đối, số trung bình và các chỉ tiêu đo độ biến thiên tiêu thức.

4.1. Số tuyệt đối trong thống kê

4.1.1. Khái niệm và ý nghĩa của số tuyệt đối

Số tuyệt đối trong thống kê biểu hiện quy mô, khối lượng của hiện tượng trong điều kiện thời gian và địa điểm cụ thể.

Số tuyệt đối có thể biểu hiện số đơn vị của tổng thể hay bộ phận (như số công nhân, số học sinh, số doanh nghiệp, số công ty,...) hoặc các trị số của một tiêu thức nào đó (như doanh thu, chi phí sản xuất, tổng sản phẩm trong nước, giá trị sản xuất công nghiệp, giá trị sản xuất nông nghiệp, ...).

Khác với các đại lượng tuyệt đối trong toán học, mỗi số tuyệt đối trong thống kê bao hàm một nội dung kinh tế - xã hội nhất định tồn tại trong một thời gian và không gian cụ thể có được qua điều tra, thu thập, tính toán, tổng hợp. Số tuyệt đối trong thống kê có đơn vị tính toán cụ thể, gắn liền với một hiện tượng trong điều kiện lịch sử nhất định. Đơn vị

tính của số tuyệt đối có thể là đơn vị tự nhiên, đơn vị giá trị hoặc đơn vị thời gian lao động:

- Đơn vị tự nhiên là đơn vị tính toán phù hợp với đặc điểm vật lý của hiện tượng, các hiện tượng này có thể được tính theo chiều dài (m, km,...) theo diện tích (m^2 , km^2 , ha...) theo trọng lượng (kg, tấn, tạ, yến,...) theo dung tích (lít, m^3 , ...). Đơn vị tự nhiên cũng có thể là số đơn vị tổng thể (cái, con, chiếc, số người, số sự kiện, số trường hợp,...). Trong nhiều trường hợp phải dùng đơn vị kép để tính toán như sản lượng điện tính bằng kw/h; khối lượng vận chuyển tính bằng tấn, km,... Trong sản xuất, những sản phẩm giống nhau về giá trị sử dụng nhưng khác nhau về kích thước, trọng lượng, công suất v.v... phải dùng đơn vị hiện vật quy ước để quy đổi các sản phẩm hiện vật về cùng một tiêu chuẩn nào đó như: Nước mắm quy theo độ đậm, xà phòng quy theo hàm lượng chất béo, than quy theo hàm lượng cacbon,...

- Đơn vị thời gian lao động như giờ - công, ngày - công,... thường được dùng để tính lượng lao động hao phí để sản xuất ra những sản phẩm không thể tổng hợp hoặc so sánh được với nhau bằng các đơn vị tính toán khác nhau hoặc những sản phẩm phức tạp do nhiều người cùng thực hiện qua nhiều giai đoạn khác nhau.

- Đơn vị giá trị (tiền tệ) như đồng Việt Nam và đô la Mỹ được sử dụng phổ biến nhất nó có thể giúp cho việc tổng hợp, so sánh nhiều sản phẩm có giá trị sử dụng và đơn vị đo lường khác nhau và dùng để so sánh quốc tế.

Số tuyệt đối có ý nghĩa quan trọng trong mọi công tác nghiên cứu kinh tế. Thông qua các số tuyệt đối ta sẽ có một nhận thức cụ thể về quy mô, khối lượng thực tế của hiện tượng nghiên cứu, nhờ các số tuyệt đối có thể biết cụ thể nguồn tài nguyên của đất nước, khả năng tiềm tàng trong nền kinh tế quốc dân, tính toán các mặt cân đối lớn, nghiên cứu các mối quan hệ, đánh giá thực hiện các mục tiêu phát triển kinh tế - xã hội của đất nước hay của mỗi địa phương, từng khu vực...

Số tuyệt đối là cơ sở đầu tiên để tiến hành phân tích thống kê đồng thời là cơ sở để tính số tương đối và số bình quân.

Số tuyệt đối là căn cứ không thể thiếu được trong việc xây dựng các kế hoạch kinh tế quốc dân và chỉ đạo thực hiện kế hoạch.

Số tuyệt đối chính xác là sự thật khách quan, có sức thuyết phục không thể phủ nhận được.

4.1.2. Các loại số tuyệt đối

Có hai loại số tuyệt đối là số tuyệt đối thời kỳ và số tuyệt đối thời điểm.

* *Số tuyệt đối thời kỳ*: Phản ánh quy mô, khối lượng của hiện tượng trong một thời kỳ nhất định. Ví dụ: Doanh số bán hàng của công ty, giá trị sản xuất công nghiệp của doanh nghiệp. Số tuyệt đối thời kỳ được hình thành thông qua sự tích lũy về lượng của hiện tượng trong suốt thời gian nghiên cứu, cho nên trị số của số tuyệt đối thời kỳ tùy thuộc vào độ dài của thời kỳ nghiên cứu. Thời kỳ nghiên cứu càng dài thì trị số của số tuyệt đối thời kỳ càng lớn và ngược lại (Ví dụ: Doanh số bán hàng 3 tháng cộng lại bằng doanh số bán hàng 1 quý, doanh số bán hàng 4 quý cộng lại bằng doanh số bán hàng cả năm).

* *Số tuyệt đối thời điểm*: Phản ánh quy mô, khối lượng của hiện tượng nghiên cứu tại một thời điểm nhất định. Ví dụ: Dân số của Việt Nam tại thời điểm 0h ngày 01/04/2009, giá trị tài sản cố định của doanh nghiệp có đến 31/12/2013,...

Số tuyệt đối thời điểm ít phụ thuộc vào độ dài của thời kỳ nghiên cứu và không thể cộng các số tuyệt đối thời điểm lại với nhau để được một số tuyệt đối thời điểm chung. Số tuyệt đối thời điểm cần được lựa chọn vào thời điểm hợp lý để đảm bảo yêu cầu nghiên cứu và so sánh.

4.2. Số tương đối

4.2.1. Khái niệm và ý nghĩa của số tương đối

Số tương đối trong thống kê biểu hiện quan hệ so sánh giữa hai mức độ của hiện tượng nghiên cứu. Đó là kết quả của việc so sánh giữa hai mức độ cùng loại nhưng khác nhau về thời gian hay không gian hoặc giữa hai mức độ khác loại nhưng có liên quan với nhau.

Số tương đối trong thống kê có thể biểu hiện bằng số lần hoặc phần trăm (%) hay đơn vị kép (người/km² ; 1000 đ/ người; ...).

Ví dụ: Giá trị sản xuất nông nghiệp của tỉnh A năm 2012 so với 2011 bằng 112,5% (tăng 12,5%) so với kế hoạch đạt 105,8% (vượt kế hoạch 5,8%). Mật độ dân số của Việt Nam năm 2011 là 265 (người/ km²).

Trong phân tích thống kê, các số tương đối được sử dụng rộng rãi để nêu lên kết cấu, quan hệ so sánh, tốc độ phát triển, trình độ phổ biến... của hiện tượng nghiên cứu trong điều kiện thời gian và không gian cụ thể. Các số tương đối không phải là con số thu thập được qua điều tra mà là kết quả so sánh giữa hai chỉ tiêu thống kê, do đó số tương đối bao giờ cũng có gốc so sánh. Tùy theo mục đích nghiên cứu, gốc dùng để so sánh có thể khác nhau. Để biểu hiện sự phát triển của hiện tượng theo thời gian thì gốc là mức độ kỳ trước, để kiểm tra tình hình thực hiện kế hoạch thì gốc là mức độ ở kỳ kế hoạch, để biểu hiện mối quan hệ giữa bộ phận với tổng thể thì gốc là mức độ tổng thể v.v... Như vậy, do sử dụng gốc so sánh khác nhau sẽ tính được nhiều số tương đối khác nhau.

Số tương đối trong thống kê có ý nghĩa quan trọng, cũng như các số tuyệt đối, số tương đối nói lên mặt lượng trong mối liên hệ mật thiết với mặt chất của hiện tượng nghiên cứu. Tuy nhiên, trong khi các số tuyệt đối mới chỉ khái quát được về quy mô, khối lượng thực tế của hiện tượng thì số tương đối cho phân tích các đặc điểm của hiện tượng, nghiên cứu các hiện tượng đó trong mối quan hệ so sánh với nhau. Trong công tác lập kế hoạch và kiểm tra tình hình thực hiện kế hoạch số tương đối đóng vai trò quan trọng.

Trong trường hợp cần giữ bí mật số tuyệt đối, người ta có thể sử dụng số tương đối để biểu hiện tình hình thực tế của hiện tượng.

4.2.2. Các loại số tương đối

Căn cứ vào nội dung mà số tương đối phản ánh, có các loại số tương đối sau đây:

4.2.2.1. Số tương đối động thái

Số tương đối động thái biểu hiện sự biến động về mức độ của hiện tượng nghiên cứu theo thời gian. Số tương đối động thái được tính bằng

cách so sánh hai mức độ cùng loại của hiện tượng ở hai thời kỳ (hay thời điểm) khác nhau. Mức độ thời kỳ được tiến hành nghiên cứu thường gọi là mức độ kỳ báo cáo, còn mức độ của một thời kỳ nào đó được chọn làm cơ sở so sánh thường gọi là mức độ kỳ gốc. Đơn vị tính của số tương đối động thái là số lần hay số phần trăm.

Công thức tính:
$$t_d = \frac{y_1}{y_0} \tag{4.1}$$

Trong đó: t_d : Số tương đối động thái.

y_1 : Mức độ kỳ nghiên cứu (kỳ báo cáo).

y_0 : Mức độ kỳ gốc.

Ví dụ: Vốn đầu tư xây dựng của một địa phương năm 2011 là 2000 tỷ đồng, năm 2012 là 2500 tỷ đồng. Nếu so sánh vốn đầu tư xây dựng năm 2012 với năm 2011 ta sẽ có số tương đối động thái (áp dụng công thức 4.1).

$$t_d = \frac{2500}{2000} = 1,25 \text{ hay } 125\%$$

Gốc so sánh của số tương đối động thái là gốc liên hoàn hoặc gốc cố định. Số tương đối động thái còn được gọi là tốc độ phát triển hay chỉ số phát triển. Muốn tính số tương đối động thái chính xác cần chú ý đảm bảo tính chất so sánh được giữa các mức độ kỳ nghiên cứu với kỳ gốc hay nói cách khác cần phải đảm bảo giống nhau về nội dung kinh tế, về phương pháp tính, về đơn vị tính toán, về phạm vi, độ dài thời gian mà mức độ phản ánh.

4.2.2.2. Số tương đối kế hoạch

Số tương đối kế hoạch là chỉ tiêu tương đối phản ánh mức cần đạt tới trong kỳ kế hoạch và mức thực tế đạt được so với kế hoạch được giao về một chỉ tiêu kinh tế - xã hội nào đó. Số tương đối kế hoạch có đơn vị tính là số lần hoặc phần trăm (%).

Số tương đối kế hoạch dùng để lập kế hoạch và đánh giá tình hình thực hiện kế hoạch về các chỉ tiêu kinh tế - xã hội.

Có hai loại số tương đối kế hoạch: Số tương đối nhiệm vụ kế hoạch và số tương đối thực hiện kế hoạch.

* Số tương đối nhiệm vụ kế hoạch: Biểu hiện quan hệ tỷ lệ giữa mức độ kỳ kế hoạch (tức là mức độ cần đạt tới của chỉ tiêu nào đó trong kỳ kế hoạch) với mức độ đã thực hiện ở kỳ gốc của chỉ tiêu nghiên cứu.

Công thức tính:
$$t_{NK} = \frac{y_{kh}}{y_0} \quad (4.2)$$

Trong đó: t_{NK} : Số tương đối nhiệm vụ kế hoạch.

y_{kh} : Mức độ kỳ kế hoạch.

y_0 : Mức độ kỳ gốc.

* Số tương đối hoàn thành kế hoạch: Biểu hiện quan hệ so sánh giữa mức độ thực tế kỳ nghiên cứu với mức độ kỳ kế hoạch trong kỳ của chỉ tiêu nghiên cứu.

Công thức tính:
$$t_{HT} = \frac{y_1}{y_{kh}} \quad (4.3)$$

Trong đó: t_{HT} : Số tương đối hoàn thành kế hoạch.

y_{kh} : Mức độ kỳ kế hoạch.

y_1 : Mức độ kỳ nghiên cứu.

Thay số vào ví dụ trên về vốn đầu tư xây dựng của một địa phương. Biết thêm kế hoạch vốn đầu tư xây dựng năm 2012 của địa phương này là 2100 tỷ đồng; áp dụng công thức 4.2 tính được số tương đối nhiệm vụ kế hoạch:

$$t_{nv} = \frac{2100}{2000} = 1,05 \text{ hay } 105\%$$

Áp dụng công thức 4.3 ta tính được số tương đối hoàn thành kế hoạch:

$$t_{HT} = \frac{2500}{2100} = 1,1904 \text{ hay } 119,04\%$$

Như vậy, nhiệm vụ kế hoạch vốn đầu tư xây dựng của địa phương năm 2012 xây dựng bằng 105% so với năm 2011 và kết quả thực hiện năm 2012 so với kế hoạch năm 2012 đạt 119,04% tức là vượt mức kế hoạch 19,04%.

Đối với những chỉ tiêu kinh tế mà kế hoạch dự kiến tăng là chiều hướng tốt thì số tương đối hoàn thành kế hoạch tính ra trên 100% là vượt mức, dưới 100% là không hoàn thành kế hoạch. Ngược lại, những chỉ tiêu kinh tế mà kế hoạch dự kiến giảm là chiều hướng tốt (giá cả, giá thành, mức tiêu hao nguyên vật liệu...) thì số tương đối hoàn thành kế hoạch tính ra dưới 100% là vượt kế hoạch, trên 100% là không hoàn thành kế hoạch.

Khi tính các số tương đối kế hoạch phải đảm bảo tính chất so sánh được về nội dung, phương pháp tính... giữa mức độ kỳ thực tế và mức độ kỳ kế hoạch.

Số tương đối kế hoạch và số tương đối động thái có mối quan hệ tích số với nhau:

$$t_d = t_{NK} \cdot t_{HT} \quad (4.4)$$

Số tương đối động thái bằng tích của số tương đối nhiệm vụ kế hoạch và số tương đối hoàn thành kế hoạch.

Theo số liệu trên ta có:

$$\frac{2500}{2000} = \frac{2100}{2000} \times \frac{2500}{2100} \text{ tức là } 1,25 = 1,05 \times 1,1904$$

Trong trường hợp thiếu tài liệu tính toán, có thể vận dụng mối quan hệ trên để tính một chỉ tiêu tương đối khi đã biết hai chỉ tiêu kia.

4.2.2.3. Số tương đối kết cấu

Số tương đối kết cấu biểu hiện tỷ trọng của mỗi bộ phận cấu thành trong một tổng thể. Số tương đối kết cấu là kết quả so sánh trị số của từng bộ phận với trị số của cả tổng thể.

Công thức tính:
$$d_i = \frac{y_b}{y_T} \quad (4.5)$$

Trong đó: d_i : Số tương đối kết cấu.

y_b : Trị số tuyệt đối từng bộ phận.

y_T : Trị số tuyệt đối của tổng thể.

Số tương đối kết cấu thường biểu hiện bằng số lần hoặc phần trăm.

Ví dụ: Giá trị sản xuất nông nghiệp của tỉnh A năm 2012 là 2000 tỷ đồng trong đó giá trị ngành trồng trọt là 1500 tỷ đồng và ngành chăn nuôi là 500 tỷ đồng. Tính các số tương đối kết cấu (Áp dụng công thức 4.5).

Tỷ trọng giá trị sản xuất ngành trồng trọt:

$$d_{TT} = \frac{1500}{2000} = 0,75 \text{ hay } 75\%$$

Tỷ trọng giá trị sản xuất ngành chăn nuôi:

$$d_{CN} = \frac{500}{2000} = 0,25 \text{ hay } 25\%$$

Muốn tính các số tương đối kết cấu được chính xác trước hết phải phân biệt các bộ phận có tính chất khác nhau trong tổng thể nghiên cứu có nghĩa phải phân tổ chính xác. Vì vậy, việc tính số tương đối kết cấu có quan hệ mật thiết với phương pháp phân tổ thống kê.

4.2.2.4. Số tương đối cường độ

Số tương đối cường độ biểu hiện trình độ phổ biến của hiện tượng trong điều kiện lịch sử nhất định. Số tương đối cường độ được tính bằng cách so sánh mức độ của hai hiện tượng khác nhau nhưng có quan hệ với

nhau. Số tương đối cường độ có đơn vị tính kép do đơn vị tính ở tử số và mẫu số của số tương đối hợp thành. Số tương đối cường độ được sử dụng rộng rãi để biểu hiện trình độ phát triển sản xuất, trình độ đảm bảo vật chất và văn hóa của nhân dân một nước hoặc so sánh sự phát triển kinh tế giữa các nước trong khu vực và trên thế giới. Các số tương đối cường độ trong thống kê thường gặp như mật độ dân số tính bằng số người dân (người) chia cho diện tích tự nhiên (km^2) với đơn vị tính là người/ km^2 ; GDP bình quân đầu người bằng GDP (1000đ) chia cho dân số trung bình (người) với đơn vị tính là 1000đ/người; sản lượng lương thực bình quân đầu người bằng (kg) chia cho dân số trung bình (người) với đơn vị tính là kg/người; Bác sĩ tính trên vạn dân bằng số bác sĩ (người) chia cho dân số (vạn người) với đơn vị tính là người/vạn người.

4.2.2.5. Số tương đối so sánh

Số tương đối so sánh biểu hiện quan hệ so sánh về mức độ giữa hai mức độ trong một tổng thể hoặc giữa hai mức độ của hai hiện tượng cùng loại nhưng khác nhau về không gian.

Công thức tính:
$$t_s = \frac{y_A}{y_B} \text{ hoặc } t_s = \frac{y_B}{y_A} \quad (4.6)$$

Trong đó: t_s : Số tương đối so sánh.

y_A : Mức độ của hiện tượng ở địa điểm A.

y_B : Mức độ của hiện tượng ở địa điểm B.

Tùy theo mục đích nghiên cứu và đặc điểm của hiện tượng có thể thay đổi mức so sánh (đảo ngược giữa tử số và mẫu số).

Ví dụ: Dân số trung bình của Việt Nam năm 2012 khu vực thành thị là 28.356,4 nghìn người. Khu vực nông thôn là 64.416,5 nghìn người. Áp dụng công thức (4.5) tính được:

- Dân số thành thị so với nông thôn:

$$t_s = \frac{28.356,4}{64.416,5} = 0,4402 \text{ hay } 44,02\%$$

- Dân số nông thôn so với thành thị:

$$t_s = \frac{64.416,5}{28.356,4} = 2,2714 \text{ hay } 227,14\%$$

Khi tính số tương đối không gian, cần chú ý tính chất có thể so sánh được giữa các chỉ tiêu.

4.2.3. Điều kiện vận dụng số tuyệt đối và số tương đối

4.2.3.1. Phải xét đến đặc điểm của hiện tượng nghiên cứu

Khi sử dụng số tuyệt đối và số tương đối phải chú ý đến đặc điểm của hiện tượng trong điều kiện thời gian và không gian khác nhau. Có những trường hợp cùng một biểu hiện về lượng, nhưng có thể mang ý nghĩa khác nhau.

Ví dụ: Một doanh nghiệp thương mại - dịch vụ có số lao động là 500 người với một doanh nghiệp sản xuất công nghiệp có 500 công nhân ta kết luận quy mô kinh doanh của doanh nghiệp thương mại với số lao động quá lớn, trong khi doanh nghiệp sản xuất công nghiệp với số công nhân 500 người, quy mô nhỏ vì đặc điểm của hai doanh nghiệp trên khác nhau.

4.2.3.2. Phải vận dụng kết hợp các số tương đối với số tuyệt đối

Số tương đối thường là kết quả so sánh của hai số tuyệt đối, số tương đối tính ra rất khác nhau, tùy thuộc vào việc lựa chọn gốc so sánh. Có khi số tương đối tính ra rất lớn nhưng ý nghĩa của nó lại không lớn, vì số tuyệt đối tương ứng với nó rất nhỏ; ngược lại có khi số tương đối tính ra nhỏ, nhưng lại có ý nghĩa quan trọng vì số tuyệt đối ứng với nó có quy mô đáng kể. Vì vậy, khi nghiên cứu thống kê nếu vận dụng kết hợp cả số tương đối và số tuyệt đối thì sẽ nhận thức được sâu sắc và chính xác đặc điểm của hiện tượng cả về quy mô, mức độ hơn kém, to nhỏ, nhanh chậm, tốc độ tăng (giảm), trình độ phổ biến...

4.3. Số trung bình trong thống kê

4.3.1. Khái niệm, ý nghĩa của số trung bình

Số trung bình trong thống kê biểu hiện mức độ đại biểu (điển hình) theo một tiêu thức nào đó của một tổng thể bao gồm nhiều đơn vị cùng loại.

Ví dụ: Thu nhập trung bình của một nhân viên trong công ty thương mại là 7,5 triệu đồng mỗi tháng.

Số trung bình có tính chất tổng hợp và khái quát cao. Chỉ dùng một trị số để nêu lên mức độ chung nhất, phổ biến nhất, có tính chất đại biểu nhất của tiêu thức nghiên cứu, không kể đến sự chênh lệch thực tế giữa các đơn vị tổng thể. Số trung bình không biểu hiện một mức độ cá biệt mà là mức độ tính chung cho các đơn vị tổng thể (thu nhập trung bình mỗi công nhân, năng suất lao động trung bình mỗi công nhân, giá thành trung bình mỗi đơn vị sản phẩm,...).

Do số trung bình chỉ biểu hiện đặc điểm chung của cả tổng thể nghiên cứu, cho nên các nét riêng biệt có tính chất ngẫu nhiên của từng đơn vị cá biệt bị loại trừ, có nghĩa là số trung bình có đặc điểm san bằng mọi chênh lệch giữa các đơn vị về trị số của tiêu thức nghiên cứu nhưng sự san bằng này chỉ có ý nghĩa khi ta tính toán số trung bình từ số lượng đủ lớn.

Số trung bình có ý nghĩa quan trọng trong công tác lý luận và trong thực tiễn, nó được dùng trong công tác nghiên cứu kinh tế nhằm nêu lên đặc điểm chung của hiện tượng kinh tế - xã hội số lớn trong điều kiện thời gian và địa điểm cụ thể.

Việc sử dụng số trung bình tạo điều kiện để so sánh giữa các hiện tượng không có cùng quy mô như so sánh tiền lương trung bình và năng suất lao động trung bình của công nhân giữa hai doanh nghiệp, so sánh năng suất thu hoạch giữa hai địa phương...

Số trung bình còn được dùng để nghiên cứu các quá trình biến động qua thời gian, nhất là các quá trình sản xuất. Sự biến động của số trung bình qua thời gian cho thấy được xu hướng phát triển cơ bản của hiện tượng số lớn, tức là của đại bộ phận các đơn vị tổng thể, trong khi từng đơn vị cá biệt không cho ta thấy những điều đó.

Số trung bình không chỉ dùng trong công tác thống kê mà còn cả trong công tác kế hoạch. Rất nhiều chỉ tiêu kế hoạch được biểu hiện bằng số trung bình.

Số trung bình chiếm một vị trí quan trọng trong việc vận dụng các phương pháp phân tích thống kê: Khi phân tích sự biến động, phân tích mối liên hệ, dự đoán thống kê, điều tra chọn mẫu đều sử dụng số trung bình.

4.3.2. Các loại số trung bình

Trong thực tế có nhiều loại số trung bình, mỗi loại có công thức tính toán khác nhau, sử dụng loại nào trong thống kê phải căn cứ vào đặc điểm của hiện tượng, nguồn tài liệu sẵn có, căn cứ vào mục đích nghiên cứu, ý nghĩa kinh tế.

4.3.2.1. Số trung bình cộng

Số trung bình cộng (còn gọi là số bình quân số học) được tính bằng cách chia tổng các lượng biến (theo một tiêu thức nào đó) cho số đơn vị tổng thể, với công thức tính như sau:

* Số trung bình cộng giản đơn

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \tag{4.7}$$

Trong đó:

x_i : Các lượng biến ($i = 1, 2, \dots, n$).

$\bar{x}$: Số trung bình.

n : Số đơn vị tổng thể.

Công thức (4.7) được vận dụng để tính các mức độ trung bình của các chỉ tiêu khi tài liệu thu thập chỉ có ít, không có phân tổ.

Ví dụ: Một cửa hàng có 5 nhân viên bán hàng, doanh số bán hàng của mỗi nhân viên trong tháng lần lượt là: 150, 160, 170, 180, 185 (tr.đ). vậy doanh số trung bình cho một nhân viên trong tháng (áp dụng công thức 4.7) là:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{150 + 160 + 170 + 180 + 185}{5} = 169 \text{ (tr.đ)}$$

Doanh số bán hàng trung bình 1 nhân viên bán hàng trong tháng ở của hàng trên là 169 triệu đồng.

* Số trung bình công gia quyền

Công thức tính:

$$\bar{x} = \frac{x_1f_1 + x_2f_2 + \dots + x_nf_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum x_if_i}{\sum f_i} \tag{4.8}$$

Trong đó:

x_i : Các lượng biến (i = 1,2,...,n).

f_i : Các quyền số (tần số).

$\bar{x}$: Số trung bình.

Công thức 4.8 được áp dụng để tính bình quân các lượng biến, trong đó, mỗi lượng biến có số lần gặp khác nhau (số lần gặp còn gọi là tần số hay ở đây gọi là quyền số). Như vậy, số trung bình gia quyền được áp dụng để tính bình quân các lượng biến trong dãy số có phân tổ.

Ví dụ: Tính năng suất lao động của công nhân theo tài liệu sau:

NSLĐ (sản phẩm/người)	Số công nhân (người)
x_i	f_i
60	10
62	20
65	12
70	8
Cộng	50

Theo số liệu trên, áp dụng công thức 4.8 tính được:

$$\begin{aligned}\bar{x} &= \frac{\sum x_i f_i}{\sum f_i} = \frac{(60 \times 10) + (62 \times 20) + (65 \times 12) + (70 \times 8)}{10 + 20 + 12 + 8} \\ &= \frac{3180}{50} = 64 \text{ (sản phẩm)}\end{aligned}$$

Chú ý:

Khi tính số trung bình theo số liệu từ dãy số phân tổ và có khoảng cách tổ, ta phải tính số trung bình tổ (trị số giữa của mỗi tổ) sau đó áp dụng công thức số trung bình cộng gia quyền theo công thức 4.8.

Số trung bình tổ (trị số giữa mỗi tổ) $x_i = \frac{x_{\max i} + x_{\min i}}{2}$

Trong đó: $x_{\max i}$, $x_{\min i}$ là giới hạn trên và giới hạn dưới của khoảng cách tổ ở mỗi tổ, trị số x_i này được coi là lượng biến đại diện cho mỗi tổ.

Ví dụ: Có bảng số liệu giả định như sau:

Năng suất lao động (1000đ/người)	Trị số giữa x_i	Số công nhân (người) f_i
5000 - 6000	5500	10
6000 - 7000	6500	20
7000 - 8000	7500	35
8000 - 9000	8500	30
9000 - 10000	9500	25
Cộng		120

Áp dụng công thức (4.8):

$$\begin{aligned}\bar{x} &= \frac{\sum x_i f_i}{\sum f_i} = \frac{(5500 \times 10) + (6500 \times 20) + (7500 \times 35) + (8500 \times 30) + (9500 \times 25)}{(10 + 20 + 35 + 30 + 25)} \\ &= \frac{94000}{120} = 7833 \text{ (nghìn đồng/người)}\end{aligned}$$

- Trường hợp các khoảng cách tổ được hình thành theo các lượng biến liên tục nhưng không có giới hạn trên và dưới trùng nhau như 500 - 599,99; 600 - 699,99; 700 - 799,99; ... thì trị số giữa tính theo giới hạn dưới của 2 tổ kế tiếp nhau.

Ví dụ : $x_1 = \frac{500 + 600}{2}$; $x_2 = \frac{600 + 700}{2}$; ...

- Đối với những dãy số lượng biến có khoảng cách tổ mở (tức là tổ đầu và tổ cuối không có giới hạn dưới và giới hạn trên) thì việc tính toán trị số giữa của các tổ này phải căn cứ vào khoảng cách tổ gần chúng nhất mà tính toán cho hợp lý.

- Trường hợp tần số (f_i) cho dưới dạng tỷ trọng (số tương đối kết cấu) ta có công thức tính:

$$\bar{x} = \frac{\sum x_i d_i}{\sum d_i} = \frac{\sum x_i d_i}{100}; \text{ Trong đó: } d_i = \frac{f_i}{\sum f_i}; \quad (4.9)$$

Ví dụ: Tính năng suất lao động trung bình theo tài liệu sau:

Phân xưởng	Năng suất lao động (tấn) x_i	Tỷ trọng số công nhân (%) d_i
A	120	20
B	100	35
C	80	45
Cộng		100

Theo số liệu trên, áp dụng công thức 4.9 năng suất lao động trung bình 1 công nhân của 3 phân xưởng:

$$\bar{x} = \frac{\sum x_i d_i}{\sum d_i} = \frac{(120 \times 20) + (100 \times 35) + (80 \times 45)}{100} = 95 \text{ (tấn/người)}$$

Vậy năng suất lao động trung bình một công nhân là 95 tấn.

4.3.2.2. Số trung bình điều hòa

Số trung bình điều hòa cũng có nội dung kinh tế như số trung bình cộng, tính được bằng cách đem chia tổng các lượng biến của tiêu thức cho số đơn vị tổng thể.

* Số trung bình điều hòa gia quyền

Nếu tài liệu chỉ có lượng biến (x_i) và tổng lượng biến, thiếu số liệu về đơn vị tổng thể, ta áp dụng công thức số trung bình điều hòa gia quyền.

$$\bar{x} = \frac{\sum M_i}{\sum \frac{M_i}{x_i}} \tag{4.10}$$

Trong đó: x_i : Các lượng biến ($i = \overline{1, n}$).

M_i : Tổng các lượng biến của tiêu thức ($i = \overline{1, n}$).

$\bar{x}$: Số trung bình.

Ví dụ: Một phân xưởng có 3 tổ sản xuất, chi phí sản xuất và giá thành đơn vị sản phẩm của các tổ như sau:

Tổ sản xuất	Chi phí sản xuất (1000đ) M_i	Giá thành đơn vị sản phẩm (1000đ) x_i	Khối lượng sản phẩm (chiếc) $\frac{M_i}{x_i}$
I	110.000	110	1000
II	120.000	100	1200
III	142.500	95	1500
Cộng	372.500		3700

Áp dụng công thức trung bình điều hòa gia quyền:

$$\bar{x} = \frac{\sum M_i}{\sum \frac{M_i}{x_i}} = \frac{372500}{3700} = 100,67 \text{ (nghìn đồng)}$$

Vậy giá thành trung bình 1 sản phẩm của 3 phân xưởng là 100,6 (ng.đ).

Nếu (M_i) cho dưới dạng tỷ trọng (số tương đối kết cấu) ta dùng công thức:

$$\bar{x} = \frac{\sum d_i}{\sum \frac{d_i}{x_i}} \quad (4.11)$$

Trong đó: $d_i = \frac{M_i}{\sum M_i}$ là tỷ trọng của tổng lượng biến.

Ví dụ: Giả sử có tỷ trọng chi phí sản xuất của tổ I là 30%; tổ II là 30%; tổ III là 40%. Giá thành đơn vị sản phẩm của tổ I là 110 (ng.đ), tổ II là 100 (ng.đ); tổ III là 95 (ng.đ). Tính giá thành trung bình 1 sản phẩm của 3 phân xưởng trên.

Áp dụng công thức:

$$\bar{x} = \frac{\sum d_i \cdot \frac{30+30+40}{110 + \frac{30}{100} + \frac{40}{95}}}{\sum \frac{d_i}{x_i}} = \frac{100}{0,273 + 0,3 + 0,421} = 100,6 \text{ (nghìn đồng)}$$

* Số trung bình điều hòa giản đơn

Trường hợp các quyền số (M_i) bằng nhau, tức là khi:

$$M_1 = M_2 = \dots = M_n = M$$

Công thức (4.9) có thể thay đổi như sau:

$$\bar{x} = \frac{\sum M_i}{\sum \frac{M_i}{x_i}} = \frac{nM}{M \sum \frac{1}{x_i}} = \frac{n}{\sum \frac{1}{x_i}} \quad (4.12)$$

Công thức (4.12) được gọi là số trung bình điều hòa giản đơn.

Ví dụ: Một nhóm 3 công nhân cùng sản xuất với thời gian lao động như nhau, người thứ nhất sản xuất 1 sản phẩm hết 15 phút, người thứ hai hết 20 phút, người thứ 3 hết 30 phút. Hãy tính thời gian lao động trung bình để sản xuất ra 1 sản phẩm của 3 công nhân nói trên?

$$\text{Áp dụng công thức (4.12) : } \bar{x} = \frac{n}{\sum \frac{1}{x_i}} = \frac{3}{\frac{1}{15} + \frac{1}{20} + \frac{1}{30}} = 20 \text{ (phút)}$$

4.3.2.3. Số trung bình nhân

Khi các lượng biến trong tổng thể có quan hệ tích số với nhau người ta áp dụng công thức số trung bình nhân.

* Số trung bình nhân giản đơn

Công thức tính:

$$\bar{x} = \sqrt[n]{x_1 x_2 \dots x_n} = \sqrt[n]{\prod x_i} \quad (4.13)$$

Trong đó: $\bar{x}$: Số trung bình

x_i : Các lượng biến ($i = \overline{1, n}$)

$\prod$: Là ký hiệu của tích

Ví dụ: Tốc độ phát triển sản xuất của 1 doanh nghiệp như sau:

Năm 2007 so với năm 2006 bằng 116%;

Năm 2008 so với năm 2007 bằng 111%;

Năm 2009 so với năm 2008 bằng 112%;

Năm 2010 so với năm 2009 bằng 113%;

Năm 2011 so với năm 2010 bằng 112%;

Năm 2012 so với năm 2011 bằng 111%.

Tính tốc độ phát triển trung bình một năm trong khoảng thời gian từ 2006 đến 2012 về tình hình sản xuất của doanh nghiệp trên?

Theo số liệu trên, áp dụng công thức (4.13) ta tính được:

$$\bar{x} = \sqrt[6]{\prod x_i} = \sqrt[6]{1,16 \times 1,11 \times 1,12 \times 1,13 \times 1,12 \times 1,11} = 1,125 \text{ hay } 112,5\%$$

Trung bình năm thời kỳ 2007 - 2012 tốc độ phát triển của doanh nghiệp là 1,125 lần hay 112,5%.

* Số trung bình nhân gia quyền

Khi các lượng biến (x_i) có các tần số (f_i) khác nhau ta có công thức số trung bình nhân gia quyền.

Công thức tính:

$$\bar{x} = \sqrt[f_i]{x_1^{f_1} x_2^{f_2} \dots x_n^{f_n}} = \sqrt[f_i]{\prod x_i^{f_i}} \quad (i = \overline{1, n}) \quad (4.14)$$

Trong đó: f_i ($i=1,2,\dots,k$) là quyền số của các lượng biến.

Ví dụ: Trong 10 năm tốc độ phát triển sản xuất ở 1 doanh nghiệp như sau: 5 năm đầu tốc độ phát triển mỗi năm là 110%, có 2 năm tiếp theo tốc độ phát triển mỗi năm là 125% và 3 năm cuối tốc độ phát triển mỗi năm là 115%. Hãy tính tốc độ phát triển trung bình hàng năm về sản xuất của doanh nghiệp trong thời kỳ 10 năm.

Áp dụng công thức (4.14):

$$\bar{x} = \sqrt[f_i]{\prod x_i^{f_i}} = \sqrt[5+2+3]{1,1^5 \times 1,25^2 \times 1,15^3} = 1,144 \text{ (lần) hay } 114,4\%$$

Tốc độ phát triển sản xuất trung bình mỗi năm thời kỳ 10 năm của doanh nghiệp là 1,144 hay 114,4%.

4.3.2.4. Một (ký hiệu là M_0)

Một là biểu hiện của tiêu thức gặp nhiều nhất trong một tổng thể hay trong một dãy số phân phối. Nó không san bằng, bù trừ chênh lệch giữa các lượng biến cho nên có thể dùng một để bổ sung hoặc thay thế cho số trung bình trong trường hợp tính số trung bình có khó khăn. Chẳng hạn, khi tính giá cả trung bình Một loại hàng hóa nào đó trên thị trường không thể đăng ký đầy đủ giá cả của tất cả các loại hàng hóa đó rồi tính trung bình mà chỉ đăng ký giá phổ biến của mặt hàng trong thời gian nghiên cứu. Cách tính Một cho dãy số lượng biến như sau:

- Đối với dãy số lượng biến không có khoảng cách tổ, Một là lượng biến có tần số lớn nhất.

- Đối với dãy số lượng biến có khoảng cách tổ:

+ Trường hợp phân tổ có khoảng cách tổ đều nhau. Tổ có chứa Một là tổ có tần số lớn nhất và Một được tính bằng công thức:

$$M_0 = x_{\min} + h \frac{(f_2 - f_1)}{(f_2 - f_1) + (f_2 - f_3)} \quad (4.15)$$

Trong đó:

$x_{\min}$: Giới hạn dưới của tổ có Một.

h : Trị số khoảng cách tổ có Một.

f_2 : Tần số của tổ có Một.

f_1, f_3 : Tần số của tổ đứng liền trước và liền sau tổ có Một.

+ Trường hợp phân tổ có khoảng cách tổ không đều nhau, tổ có Một là tổ có mật độ phân phối lớn nhất. Sau đó ta áp dụng công thức sau đây để tính giá trị của M_0 :

$$M_0 = x_{\min} + h \frac{(m_2 - m_1)}{(m_2 - m_1) + (m_2 - m_3)} \quad (4.16)$$

Trong đó:

$x_{\min}$: Giới hạn dưới của tổ có Một.

h : Trị số khoảng cách tổ có Một.

m_2 : Mật độ phân phối của tổ có Một.

m_1, m_3 : Mật độ phân phối của tổ đứng liền trước và liền sau tổ có Một.

4.3.2.5. Số trung vị (ký hiệu M_e)

Số trung vị là lượng biến của đơn vị tổng thể đứng vị trí giữa trong tổng số các đơn vị của dãy số lượng biến.

- Đối với dãy số lượng biến không có khoảng cách tổ thì số trung vị được tính cụ thể như sau:

- Nếu số đơn vị tổng thể lẻ, thì số trung vị là lượng biến của đơn vị tổng thể đứng ở chính giữa của dãy số lượng biến:

Ta có:
$$M_e = x_m \quad (4.17)$$

x_m : Lượng biến của đơn vị tổng thể thứ m ở chính giữa của dãy số lượng biến.

- Nếu số đơn vị tổng thể chẵn, thì tổng thể đó có hai đơn vị đứng ở giữa, do đó giá trị trung vị là số trung bình cộng của hai lượng biến ở giữa.

$$M_e = \frac{x_m + x_{m+1}}{2} \quad (4.18)$$

Trong đó:

x_m : Lượng biến của đơn vị tổng thể đứng giữa thứ nhất.

x_{m+1} : Lượng biến của đơn vị tổng thể đứng giữa thứ hai.

- Đối với một dãy số lượng biến có khoảng cách tổ trước hết cần xác định tổ chứa số trung vị. Muốn tìm tổ chứa số trung vị, ta cộng dồn tần số các tổ, bắt đầu cộng từ tổ thứ nhất, thứ hai,... khi nào tổng các tần số cộng được lớn hơn hoặc bằng một nửa số đơn vị tổng thể thì dừng lại và xác định tổ có số trung vị, sau đó tính trị số trung vị theo công thức:

$$M_e = x_{\min} + h \frac{\frac{\sum f_i}{2} - S}{f_{Me}} \quad (4.19)$$

Trong đó:

$x_{\min}$: Giới hạn dưới của tổ có trung vị.

h : Trị số khoảng cách tổ có trung vị.

f_{Me} : Tần số của tổ có trung vị.

S : Tổng tần số của những tổ đứng trước tổ có trung vị.

Ví dụ: Có tài liệu về năng suất lao động ở một DN như sau:

Tổ	Mức lương (1000đ/người)	Số CN (người)	Tần số tích lũy
1	2000 - 3000	35	35
2	3000 - 4000	70	105
3	4000 - 5000	95	200
4	5000 - 6000	100	300
5	6000 - 7000	60	360
6	7000 - 8000	20	380
	Cộng	380	

Tính Mốt và số trung vị về tiền lương của công nhân ở DN trên?

Theo số liệu trên, tổ 4 là tổ có Mốt, áp dụng công thức 4.15 ta tính được M_0 :

$$M_0 = 5000 + 1000 \times \frac{(100 - 95)}{(100 - 95) + (100 - 60)} = 5111,11 \text{ (Nghìn đồng/người)}$$

Mốt về tiền lương của cá nhân trong doanh nghiệp là 5111,11 nghìn đồng/người.

Từ số liệu đã cho, xác định được tổ 3 là tổ có số trung vị và áp dụng công thức 4.19 ta tính được:

$$M_e = 4000 + 1000 \times \frac{(190 - 105)}{95} = 4894,74 \text{ (Nghìn đồng/người)}$$

Số trung vị về tiền lương của doanh nghiệp là 4894,74 nghìn đồng/người.

Số trung vị có ưu điểm là không chịu ảnh hưởng của các lượng biến đột xuất, việc tính toán nhanh gọn, được dùng để bổ sung hoặc thay thế cho số trung bình cộng khi cần thiết, song không thể dùng số trung vị để dự đoán, vì không chính xác bằng số trung bình cộng.

- Tứ phân vị và thập phân vị

Trong phân tích kinh tế - xã hội, nhiều khi phải tính đến thứ bậc của các đơn vị, nghĩa là chia các đơn vị của tổng thể nghiên cứu trong dãy số phân phối thành các phần bằng nhau: Ba, bốn, năm... mười phần rồi tính giá trị của các đơn vị đứng ở vị trí cơ bản. Tùy theo vị trí của các đơn vị trong dãy số mà có tên gọi khác nhau như: Tam phân vị, tứ phân vị,... cách tính các lượng biến thứ tự này giống cách tính số trung vị, nghĩa là tính tần số tích lũy để tìm vị trí của đơn vị cần tìm, rồi đối chiếu ra lượng biến. Nếu dãy số lượng biến có khoảng cách tổ, ta tìm tổ chứa tứ phân vị, thập phân vị. Sau đó áp dụng công thức:

Tứ phân vị

$$Q_1 = x_0 + h_{Q_1} \frac{\frac{1}{4} \sum f - S_{Q_1} - 1}{f_{Q_1}} \quad (4.20)$$

Thập phân vị:

$$Q_1 = x_0 + h_{Q_1} \frac{\frac{1}{10} \sum f - S_{Q_1} - 1}{f_{Q_1}} \quad (4.21)$$

Trong đó:

x_0 : Giới hạn đầu của tổ chứa tứ phân vị hoặc thập phân vị.

h_{Q_1} : Khoảng các tổ của tổ chứa tứ phân vị và thập phân vị.

$\sum f$: Tổng các tần số trong tổng thể.

S_{Q_1-1} : Tổng các tần số của những tổ đứng trước tổ chứa tứ phân vị, thập phân vị.

f_{Q_1} : Tần số của tổ chứa tứ phân vị, thập phân vị.

Tứ phân vị, thập phân vị thường được sử dụng khi người ta muốn biết mức đạt cao nhất của 1/10 hay 1/4 số đơn vị tổng thể (xếp từ thấp lên) hoặc mức tối thiểu của 1/10 hay 1/4 số đơn vị tổng thể (xếp từ cao xuống).

4.3.3. Điều kiện vận dụng số trung bình

Tuy số trung bình có nhiều tác dụng quan trọng đối với nghiên cứu thống kê, nhưng bản thân nó có những nhược điểm: San bằng mọi chênh lệch thực tế giữa các đơn vị cá biệt; làm cho tổng thể rất phức tạp trở thành đơn giản v.v... Đây chính là chỗ dễ bị lợi dụng trong thống kê. Vì vậy, muốn vận dụng số trung bình một cách khoa học và chính xác cần chú ý các điều kiện sau:

- Thứ nhất, số trung bình chỉ được tính ra từ tổng thể đồng chất. Tổng thể đồng chất là tổng thể bao gồm những đơn vị, hiện tượng có cùng chung một tính chất, thuộc cùng một loại hình kinh tế - xã hội theo một tiêu thức nào đó.

Trong tổng thể đồng chất, sự chênh lệch về lượng cụ thể giữa các đơn vị không lớn lắm vì bản chất của nó giống nhau. Có sự chênh lệch chỉ là do các nhân tố ngẫu nhiên tác động, khi tính số trung bình các nhân tố ngẫu nhiên sẽ bù trừ và triệt tiêu. Số trung bình sẽ thể hiện đầy đủ ý nghĩa và mức độ đại biểu cho tất cả các mức độ khác nhau trong tổng thể.

- Thứ hai, số trung bình chung cần được vận dụng kết hợp với số trung bình tổ hay dãy số phân phối.

Khi so sánh, phân tích hiện tượng, nếu chỉ xét qua số trung bình chung thì chênh lệch coi như bị san bằng, các đơn vị có mức độ cao thấp khác nhau đều bị che lấp, điều đó hạn chế tác dụng của việc phân tích thống kê. Không giải thích được hết nguyên nhân và xu hướng phát triển của hiện tượng. Thậm chí, nếu không chú ý còn rút ra kết luận sai lệch. Nhiệm vụ nghiên cứu của thống kê là đi đôi với việc tìm hiểu mức độ đại biểu chung còn phải tìm ra những đơn vị điển hình, tiên tiến, đơn vị lạc hậu. Phát hiện những mầm mống mới phát sinh hoặc những bộ phận lạc hậu đang kìm hãm sự phát triển chung.

Vì những lý do trên, khi phân tích thống kê, không thể chỉ thỏa mãn với con số trung bình chung, mà còn cần bổ sung bằng các số trung bình tổ (hoặc dãy số phân phối).

4.4. Nghiên cứu độ biến thiên của tiêu thức

4.4.1. Ý nghĩa nghiên cứu độ biến thiên của tiêu thức

Số trung bình chỉ nêu lên mức độ đại biểu có tính chất chung nhất của tổng thể nghiên cứu. Mức độ này không phản ánh chênh lệch thực tế giữa các đơn vị cá biệt, có khi bản thân nội bộ hiện tượng đã có nhiều thay đổi đáng kể về mặt lượng, nhưng số trung bình tính ra không thay đổi hoặc thay đổi ít. Vì vậy, trong phân tích thống kê, không nên chỉ hạn chế ở việc tính mức độ trung bình mà cần đánh giá độ biến thiên của tiêu thức. Nghiên cứu độ biến thiên của tiêu thức có nhiều tác dụng quan trọng về lý luận và thực tiễn.

- Thứ nhất, độ biến thiên tiêu thức giúp ta đánh giá trình độ đại biểu của số trung bình. Trị số này tính ra càng lớn, độ biến thiên tiêu thức càng nhiều, do đó trình độ đại biểu của số trung bình càng thấp và ngược lại.

- Thứ hai, quan sát độ biến thiên tiêu thức trong một dãy số lượng biến sẽ thấy được nhiều đặc trưng của dãy số như: Đặc trưng về phân phối; kết cấu; tính chất đồng đều của tổng thể...

- Thứ ba, trong phân tích hoàn thành kế hoạch, độ biến thiên tiêu thức giúp ta thấy được chất lượng công tác và nhịp điệu hoàn thành kế hoạch chung, cũng như của từng bộ phận, phát hiện khả năng tiềm tàng của đơn vị.

- Thứ tư, độ biến thiên của tiêu thức còn được sử dụng trong nhiều trường hợp nghiên cứu thống kê khác như: Phân tích sự biến động, mối liên hệ, dự báo thống kê, điều tra chọn mẫu v.v...

4.4.2. Các chỉ tiêu đo độ biến thiên của tiêu thức

4.4.2.1. Khoảng biến thiên

Khoảng biến thiên là độ chênh lệch giữa lượng biến lớn nhất và lượng biến nhỏ nhất của tiêu thức nghiên cứu.

Công thức tính:
$$R = x_{\max} - x_{\min} \quad (4.22)$$

Ví dụ: Mức năng suất lao động của công nhân hai tổ sản xuất như sau:

Tổ 1: 40, 50, 60, 70, 80 (kg)

Tổ 2: 58, 59, 60, 61, 62 (kg)

Mức năng suất lao động trung bình của mỗi tổ đều là 60 kg, nhưng thực ra ở hai tổ công nhân này rất khác nhau về đồng đều theo tay nghề. Để hiểu rõ điều đó, ta cần tính khoảng biến thiên của mỗi tổ:

$$R_1 = 80 - 40 = 40 \text{ kg}$$

$$R_2 = 62 - 58 = 4 \text{ kg}$$

Kết quả cho thấy R_1 cao hơn R_2 có nghĩa là độ biến thiên của tiêu thức trong tổ 1 lớn hơn và vì thế tính chất đại biểu của số trung bình tổ 1 thấp hơn. Khoảng biến thiên càng nhỏ thì tổng thể càng đồng đều, số trung bình càng có tính chất đại biểu cao và ngược lại.

Khoảng biến thiên tính đơn giản, song chỉ phụ thuộc vào giá trị đầu và cuối; không tính đến độ lệch giữa các lượng biến khác trong dãy số.

4.4.2.2. Độ lệch tuyệt đối bình quân

Độ lệch tuyệt đối bình quân là trung bình cộng các trị số tuyệt đối của các độ lệch giữa các lượng biến và trung bình của các lượng biến.

- Tính giản đơn áp dụng cho trường hợp không phân tổ:

$$\bar{d} = \frac{\sum |x_i - \bar{x}|}{n} \quad (4.23)$$

- Tính gia quyền áp dụng cho trường hợp có phân tổ:

$$\bar{d} = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i} \quad (4.24)$$

4.4.2.3. Phương sai

Phương sai là số trung bình cộng của bình phương các độ lệch giữa lượng biến với số trung bình của các lượng biến đó.

Công thức tính:

- Tính giản đơn áp dụng cho trường hợp tài liệu không phân tổ:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \tag{4.25}$$

- Tính gia quyền áp dụng cho trường hợp tài liệu được phân tổ:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i} \tag{4.26}$$

Trong đó: f_i : Các tần số ($i = \overline{1, n}$).

Ví dụ: Có số liệu về tình hình sản xuất sản phẩm trong một ngày của công nhân một phân xưởng như sau:

Số lượng sản phẩm sản xuất trong ngày (sản phẩm) - x_i	Số công nhân (người) f_i	Phần tính toán		
		$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 . f_i$
4	2	-2	4	8
5	4	-1	1	4
6	9	0	0	0
7	3	1	1	3
8	2	2	4	8
Cộng	20	-	-	23

$$\text{Áp dụng công thức: } \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i} = \frac{23}{20} = 1,15 \text{ (sản phẩm)}$$

Công thức tính phương sai cho thấy: Sự khác nhau về dấu giữa các độ chênh lệch đã được khắc phục bằng cách tính bình phương. Song cũng chính vì thế mà trị số bị khuếch đại, đơn vị tính không phù hợp với tài liệu thực tế.

4.4.2.4. Độ lệch tiêu chuẩn

Độ lệch tiêu chuẩn là căn bậc hai của phương sai:

Công thức tính:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (4.27)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 f_i}{\sum_{i=1}^n f_i}} \quad (4.28)$$

Vẫn số liệu ở bảng ta có: $\sigma = \sqrt{\frac{23}{20}} = 1,07$

4.4.2.5. Hệ số biến thiên

Các chỉ tiêu: Độ lệch tiêu chuẩn, phương sai đều là trị số tuyệt đối có đơn vị tính. Các trị số này không những phụ thuộc vào mức độ dao động của tiêu thức, mà còn phụ thuộc vào trị số của lượng biến và số trung bình. Vì vậy, không thể dùng những chỉ tiêu này để so sánh biến thiên tiêu thức của các hiện tượng khác nhau. Hoặc giữa các hiện tượng cùng loại nhưng có số trung bình không bằng nhau. Để có thể so sánh được, ta tính tỷ số giữa độ lệch tiêu chuẩn với số trung bình của các lượng biến. Chỉ tiêu này được gọi là hệ số biến thiên.

Công thức tính:

$$V = \frac{d}{\bar{x}} \quad (4.29)$$

$$V = \frac{\sigma}{\bar{x}} \quad (4.30)$$

TÓM TẮT CHƯƠNG 4

Đặc trưng về mặt lượng của hiện tượng kinh tế - xã hội mà thống kê nghiên cứu được sử dụng phổ biến trong nghiên cứu thống kê bao gồm: Số tuyệt đối, số tương đối, số trung bình và các chỉ tiêu đo độ biến thiên tiêu thức.

Số tuyệt đối trong thống kê là chỉ tiêu biểu hiện quy mô khối lượng của hiện tượng nghiên cứu trong điều kiện thời gian và không gian cụ thể. Trong thống kê phân biệt hai loại số tuyệt đối: Số tuyệt đối thời kỳ biểu hiện mức độ tích lũy của hiện tượng nghiên cứu trong một độ dài thời gian nhất định. Số tuyệt đối thời điểm biểu hiện mức độ của hiện tượng nghiên cứu tại một thời điểm.

Số tương đối biểu hiện quan hệ so sánh giữa hai mức độ của hiện tượng nghiên cứu, tùy theo nội dung so sánh, phân biệt năm loại số tương đối: Số tương đối động thái, số tương đối kết cấu, số tương đối so sánh, số tương đối cường độ, số tương đối kế hoạch.

Số trung bình biểu hiện mức độ đại diện của một tổng thể bao gồm nhiều đơn vị cùng loại. Mức độ trung bình được tính bằng công thức trung bình cộng, trung bình điều hòa và trung bình nhân. Trong một số trường hợp có thể sử dụng Một và Trung vị để bổ sung hoặc thay thế cho số trung bình cộng.

Các chỉ tiêu đo độ biến thiên là tiêu thức sử dụng để đánh giá độ đồng đều của tổng thể nghiên cứu. Thống kê sử dụng 5 chỉ tiêu để đo độ biến thiên là: Khoảng biến thiên, độ lệch tuyệt đối trung bình, phương sai, độ lệch tiêu chuẩn, hệ số biến thiên.

CÂU HỎI ÔN TẬP CHƯƠNG 4

- 1. Khái niệm, ý nghĩa và đặc điểm của số tuyệt đối trong thống kê?
- 2. Phân biệt các loại số tuyệt đối trong thống kê?
- 3. Khái niệm, ý nghĩa và đặc điểm của số tương đối trong thống kê?
- 4. Trình bày các loại số tương đối trong thống kê?
- 5. Khái niệm, ý nghĩa và đặc điểm của số trung bình trong thống kê?
- 6. Trình bày các loại số trung bình trong thống kê?
- 7. Phân tích điều kiện vận dụng số tương đối, số tuyệt đối và số trung bình trong thống kê?
- 8. Trình bày các chỉ tiêu đo độ biến động của tiêu thức?

BÀI TẬP CHƯƠNG 4

Bài 1:

Có tài liệu về giá trị sản lượng của 1 công ty như sau:

Đơn vị tính: Triệu đồng

Phân xưởng	Quý I	Quý II	
	Thực tế	Kế hoạch	Thực tế
A	4.000	4.200	4.368
B	6.600	7.260	7.623
C	5.000	5.400	5.508

Yêu cầu tính:

- a) Các số tương đối kế hoạch của mỗi phân xưởng và toàn công ty?
- b) Số tương đối động thái mỗi phân xưởng và toàn công ty?
- c) Số tương đối kết cấu về giá trị sản lượng của từng phân xưởng.
Cho nhận xét?

Bài 2:

Năm 2014, mức bán ra của công ty X là 11.500 triệu đồng. Năm 2015 dự kiến mức bán ra tăng 25% so với năm 2014. Thực tế năm 2015 mức bán ra là 17.135 triệu đồng; số nhân viên bình quân là 100 người, trong đó số nhân viên trực tiếp kinh doanh là 65 người.

Yêu cầu tính:

- a) Số tương đối hoàn thành kế hoạch mức bán ra năm 2015?
- b) Số tương đối so sánh về số nhân viên năm 2015 của công ty trên?

Bài 3:

Có tài liệu về tình hình sản xuất của công ty X như sau:

Phân xưởng	Số công nhân (người)	Năng suất lao động bình quân mỗi công nhân (tấn)	Giá thành bình quân 1 tấn sản phẩm (trđ)
A	100	25	8,2
B	80	26	8,0
C	200	30	8,5

Yêu cầu tính:

- a) Năng suất lao động trung bình một công nhân của công ty?
- b) Giá thành trung bình mỗi tấn sản phẩm toàn công ty?

Bài 4:

Có tài liệu về sản lượng và giá thành 1 loại sản phẩm tại hai doanh nghiệp như sau:

Đợt sản phẩm	Doanh nghiệp A		Doanh nghiệp B	
	Tỷ trọng sản lượng (%)	Giá thành đơn vị SP (1000đ)	Tỷ lệ % chi phí sản xuất (%)	Giá thành đơn vị SP (1000đ)
Đợt 1	20	130	30	132
Đợt 2	35	128	60	127
Đợt 3	45	125	10	120

Hãy so sánh giá thành trung bình đơn vị sản phẩm giữa hai doanh nghiệp và rút ra nhận xét.

Bài 5:

Có tài liệu về tình hình sản xuất ở doanh nghiệp X như sau:

Phân xưởng	Quý I		Quý II	
	Giá trị sản xuất thực tế (trđ)	% hoàn thành kế hoạch	Kế hoạch giá trị sản xuất (trđ)	% hoàn thành kế hoạch
A	990	110	918	108
B	686	98	598	95
C	650	102	680	105

Hãy tính:

- a) Tỷ lệ % hoàn thành kế hoạch trung bình trong quý I; quý II; trong 6 tháng của doanh nghiệp trên?
- b) Tỷ lệ % hoàn thành kế hoạch 6 tháng của mỗi phân xưởng?
- c) Với tài liệu trên, có thể tính được những loại chỉ tiêu tương đối và chỉ tiêu trung bình nào? Hãy tính các chỉ tiêu đó?

Bài 6:

Có tài liệu về năng suất thu hoạch và sản lượng một loại cây trồng của 10 xã trong huyện T trong năm như sau:

Số thứ tự	Năng suất thu hoạch trung bình 1 ha (tấn)	Sản lượng (tấn)
1	30	10.500
2	33	19.240
3	32	18.640
4	29	19.280
5	27	19.180
6	25	16.500
7	37,8	22.852
8	38	21.780
9	36,4	19.100
10	37,5	20.500

Yêu cầu:

- a) Tính năng suất thu hoạch trung bình của toàn huyện?
- b) Nếu các xã đều đạt mức năng suất thu hoạch trung bình và trung bình trên thì sản lượng loại cây trồng này của toàn huyện sẽ tăng bao nhiêu tấn?

Bài 7:

Có tài liệu ở công ty thương mại X:

Cửa hàng	Tháng 1		Tháng 2	
	Tỷ trọng mức tiêu thụ (%)	Năng suất lao động trung bình 1 nhân viên (triệu đồng)	Kết cấu số nhân viên (%)	Năng suất lao động trung bình 1 nhân viên (triệu đồng)
A	40	100	50	120
B	30	90	30	110
C	30	80	20	100

Yêu cầu:

- a) So sánh năng suất lao động trung bình 1 nhân viên trong toàn công ty tháng 2 với tháng 1?
- b) Tính số tương đối động thái và số tương đối so sánh về năng suất lao động của các cửa hàng. Cho nhận xét?

Bài 8:

Có tài liệu về giá thành đơn vị sản phẩm ở một doanh nghiệp:

Năm 2011 so với năm 2012: 97%?

Năm 2012 so với năm 2013: 95%?

Năm 2013 so với năm 2014: 92%?

Năm 2015 so với năm 2014: 90%?

Hãy tính tốc độ phát triển trung bình một năm về giá thành đơn vị sản phẩm ở doanh nghiệp trong thời gian trên?

Bài 9:

Có tài liệu về tình hình năng suất lao động của doanh nghiệp thương mại X như sau:

Năng suất lao động (tạ/ha)	Số công nhân
35 - 40	10
40 - 45	20
45 - 50	30
50 - 60	35
60 - 80	5

Yêu cầu:

- a) Tính năng suất lao động trung bình 1 CN của doanh nghiệp?
- b) Một và trung vị về năng suất lao động. Cho nhận xét?

Bài 10:

Có tình hình thực hiện kế hoạch bán ra ở một doanh nghiệp:

Cửa hàng	Kế hoạch mức bán ra (trđ)	% hoàn thành kế hoạch bán ra
Số 1	4.000	90
Số 2	3.000	120
Số 3	2.000	100
Số 4	6.000	110
Số 5	5.000	112

Hãy tính:

- a) % hoàn thành kế hoạch bán ra trung bình toàn doanh nghiệp?
- b) Phương sai về % hoàn thành kế hoạch bán ra?
- c) Hệ số biến thiên về % hoàn thành kế hoạch?
- d) Nếu tất cả các cửa hàng đều hoàn thành và hoàn thành vượt mức kế hoạch bán ra thì tổng mức bán ra thực tế của doanh nghiệp sẽ là bao nhiêu?

Chương 5

HỒI QUY VÀ TƯƠNG QUAN

Theo quan điểm duy vật biện chứng thì các hiện tượng tồn tại trong mỗi liên hệ phổ biến, đa dạng và phong phú, giữa chúng có mối quan hệ mật thiết với nhau, tác động qua lại lẫn nhau. Không có một hiện tượng nào phát sinh, phát triển một cách tách rời cô lập với các hiện tượng khác. Vì vậy, việc nghiên cứu mối liên hệ giữa các hiện tượng là một trong những nhiệm vụ quan trọng của thống kê. Phương pháp hồi quy và phương pháp tương quan là hai phương pháp của thống kê toán được sử dụng trong phân tích mối liên hệ tương quan giữa các hiện tượng kinh tế - xã hội. Nội dung chương này trình bày việc vận dụng phương pháp hồi quy và tương quan để phân tích mối liên hệ giữa các tiêu thức nguyên nhân ảnh hưởng tới tiêu thức kết quả. Cụ thể trong các trường hợp: Mối liên hệ tương quan tuyến tính (giữa hai tiêu thức, nhiều tiêu thức) và mối liên hệ tương quan phi tuyến tính.

5.1. Mối liên hệ giữa các hiện tượng, nhiệm vụ của phương pháp hồi quy và tương quan

5.1.1. Mối liên hệ giữa các hiện tượng

Trong quá trình tồn tại và phát triển, mọi sự vật hiện tượng đều có mối liên hệ với nhau theo quy luật phụ thuộc nhân quả. Vì vậy, để hiểu một cách sâu sắc bản chất, quy luật phát triển của hiện tượng đòi hỏi thống kê phải phân tích mối liên hệ tác động tương quan giữa chúng.

Thực tế mối liên hệ giữa các hiện tượng kinh tế - xã hội rất đa dạng phong phú, tính chất và hình thức rất khác nhau. Có mối liên hệ giữa hai hiện tượng, nhiều hiện tượng hoặc giữa các bộ phận cấu thành hiện tượng phức tạp; Mức độ liên hệ phụ thuộc giữa các hiện tượng này cũng rất khác nhau. Chính vì vậy, hình thành nhiều loại mối liên hệ giữa các hiện tượng kinh tế - xã hội:

Mối liên hệ giữa các hiện tượng kinh tế - xã hội có thể diễn ra trong không gian và thời gian. *Liên hệ trong không gian* là sự tác động qua lại, sự phụ thuộc vào nhau khi chúng ở trong cùng một thời gian. *Liên hệ trong thời gian* là sự tác động qua lại, sự phụ thuộc vào nhau khi chúng ở các giai đoạn phát triển khác nhau. Do đó, khi phân tích mối liên hệ giữa các hiện tượng phải đặt chúng trong điều kiện thời gian và không gian nhất định.

Nếu xét theo mức độ liên hệ phụ thuộc giữa các hiện tượng, có thể phân biệt liên hệ hàm số và liên hệ tương quan.

Liên hệ hàm số là mối liên hệ hoàn toàn chặt chẽ giữa hai hiện tượng nghiên cứu: Có nghĩa là khi hiện tượng này thay đổi thì nó hoàn toàn quyết định đến sự thay đổi của hiện tượng có liên quan theo một tỷ lệ tương ứng. Liên hệ hàm số được viết dưới dạng khái quát $Y = f(x)$. Mối liên hệ này có thể thấy được không những ở toàn bộ tổng thể mà cả trên từng đơn vị cá biệt. Các mối liên hệ hàm số thường được biểu hiện khi nghiên cứu các hiện tượng tự nhiên như trong toán học, vật lý... nhưng hiếm gặp trong các hiện tượng kinh tế - xã hội.

Liên hệ tương quan là mối liên hệ không hoàn toàn chặt chẽ giữa các hiện tượng nghiên cứu: Khi hiện tượng này thay đổi có thể làm cho hiện tượng có liên quan thay đổi theo, nhưng không có ảnh hưởng hoàn toàn quyết định sự thay đổi đó. Các mối liên hệ này không hoàn toàn chặt chẽ, không được biểu hiện rõ trên từng đơn vị cá biệt mà phải thông qua quan sát số lớn các đơn vị. Ví dụ: Mối liên hệ giữa chi phí cho quảng cáo với doanh thu tiêu thụ sản phẩm, mối liên hệ giữa năng suất lao động với giá thành đơn vị sản phẩm trong doanh nghiệp... Khi chi phí quảng cáo sản phẩm tăng lên thì có thể làm cho doanh thu tiêu thụ sản phẩm của doanh nghiệp tăng và ngược lại, nhưng sự tăng giảm này không theo một tỷ lệ nhất định và không giống nhau ở các đơn vị. Đó là vì doanh thu của doanh nghiệp tăng hay giảm còn phụ thuộc vào nhiều nguyên nhân khác chứ không phải hoàn toàn do chi phí quảng cáo quyết định. Hay khi năng suất lao động tăng lên có thể làm cho giá thành đơn vị sản phẩm giảm xuống và ngược lại, nhưng năng suất lao động không phải là yếu tố hoàn toàn quyết định tới sự thay đổi của giá thành. Như vậy, có thể nói giữa

chi phí quảng cáo và doanh thu, giữa năng suất lao động và giá thành đơn vị sản phẩm có mối liên hệ tương quan. Mối liên hệ tương quan thường gặp khi nghiên cứu các hiện tượng kinh tế - xã hội.

Nếu xét theo chiều hướng của mối liên hệ, có thể phân biệt liên hệ thuận và liên hệ nghịch. *Liên hệ thuận* được biểu hiện khi trị số của tiêu thức nguyên nhân và trị số của tiêu thức kết quả phát triển theo cùng một hướng (cùng tăng hoặc cùng giảm). Ngược lại, *liên hệ nghịch* được biểu hiện khi trị số của tiêu thức nguyên nhân và trị số của tiêu thức kết quả phát triển ngược chiều. Ở ví dụ trên thì mối liên hệ giữa chi phí cho quảng cáo và doanh thu tiêu thụ sản phẩm là mối liên hệ thuận, giữa năng suất lao động và giá thành là mối liên hệ nghịch.

5.1.2. Nhiệm vụ của phương pháp hồi quy và tương quan

Hồi quy và tương quan là hai phương pháp của toán học, được vận dụng trong thống kê để nghiên cứu mối liên hệ tương quan giữa các hiện tượng kinh tế - xã hội. Hai phương pháp này có liên quan chặt chẽ với nhau và đều xuất phát từ cùng mục đích nghiên cứu, nên có thể gọi ngắn gọn là phương pháp tương quan.

Trong phạm vi chương này thống kê sẽ vận dụng phương pháp hồi quy và tương quan để phân tích mối liên hệ không hoàn toàn chặt chẽ giữa các hiện tượng, thể hiện thông qua mối liên hệ giữa các tiêu thức. Các tiêu thức cần nghiên cứu, sẽ được chia thành các tiêu thức nguyên nhân và các tiêu thức kết quả. Khi tiêu thức kết quả biến thiên có thể do ảnh hưởng hàng loạt của các tiêu thức nguyên nhân với mức độ ảnh hưởng khác nhau. Song, tùy theo mục đích nghiên cứu, ta có thể lựa chọn ra một hoặc một số tiêu thức nguyên nhân cơ bản, có ý nghĩa hơn cả để xác định mối liên hệ giữa chúng với giả định các tiêu thức khác không thay đổi. Những tiêu thức được chọn ra để nghiên cứu bao giờ cũng có một tiêu thức kết quả và còn lại là các tiêu thức nguyên nhân. Ví dụ, giá thành đơn vị sản phẩm có thể do ảnh hưởng của rất nhiều các nguyên nhân khác nhau như: Năng suất lao động, giá cả nguyên, nhiên, vật liệu, tiền lương và các khoản chi phí khác. Trong các doanh nghiệp sản xuất ta nhận thấy năng suất lao động có ảnh hưởng đáng kể tới giá thành đơn vị sản phẩm. Do đó, có thể chọn riêng hai tiêu thức năng suất lao động (tiêu

thức nguyên nhân) và giá thành (tiêu thức kết quả) để nghiên cứu mối liên hệ giữa chúng, còn các tiêu thức khác giả định không thay đổi.

Phương pháp hồi quy và tương quan nhằm giải quyết hai nhiệm vụ cơ bản:

- *Xác định mô hình hồi quy biểu hiện mối liên hệ*, nghĩa là xét xem mối liên hệ giữa các tiêu thức nghiên cứu được biểu hiện dưới dạng mô hình nào: Liên hệ tuyến tính (mô hình đường thẳng) hay phi tuyến tính (mô hình đường cong), liên hệ thuận hay liên hệ nghịch...

Để giải quyết nhiệm vụ này cần phải tiến hành các bước chủ yếu sau:

+ Dựa trên cơ sở phân tích lý luận để giải thích về sự tồn tại thực tế và bản chất của mối liên hệ giữa các hiện tượng nghiên cứu.

+ Kết hợp phân tích lý luận với việc thăm dò mối liên hệ đó bằng các phương pháp thống kê như: Phương pháp đồ thị, phương pháp phân tổ, phương pháp số bình quân..., hoặc dựa trên cơ sở các kết quả nghiên cứu có từ trước về hiện tượng.

+ Lựa chọn phương trình hồi quy để biểu hiện mối liên hệ. Muốn xác định đúng phương trình phải căn cứ vào số tiêu thức được chọn, hình thức và chiều hướng của mối liên hệ.

+ Tính toán và nêu ý nghĩa của các tham số của phương trình hồi quy.

- *Đánh giá trình độ chặt chẽ của mối liên hệ bằng các chỉ tiêu hệ số tương quan, tỉ số tương quan...* Từ kết quả tính các chỉ tiêu này có thể xác định vai trò ảnh hưởng của từng nguyên nhân, giải thích sự tồn tại hay không tồn tại mối liên hệ tương quan và kiểm định lại giả thiết và sự phù hợp của mô hình hồi quy đã chọn.

Trong phân tích tương quan, số tiêu thức được chọn ra nghiên cứu càng nhiều thì quá trình tính toán càng phức tạp. Dưới đây trình bày cách vận dụng cụ thể phương pháp hồi quy và tương quan trong một số trường hợp tiêu biểu nhất.

5.2. Liên hệ tương quan tuyến tính giữa hai tiêu thức số lượng

5.2.1. Phương trình hồi quy tuyến tính

Mối liên hệ tương quan giữa các hiện tượng kinh tế - xã hội được biểu hiện dưới dạng phương trình tuyến tính hoặc phi tuyến tính. Song, trong nhiều trường hợp phân tích thông thường để đơn giản quá trình tính toán, nếu với sai số cho phép người ta có thể sử dụng dạng phương trình tuyến tính để mô tả một cách gần đúng mối liên hệ mà quá trình tính toán lại đơn giản hơn.

Xét ví dụ về mối liên hệ giữa thâm niên công tác và tiền lương trong năm qua tài liệu điều tra khảo sát ngẫu nhiên 10 công nhân tại một doanh nghiệp như sau:

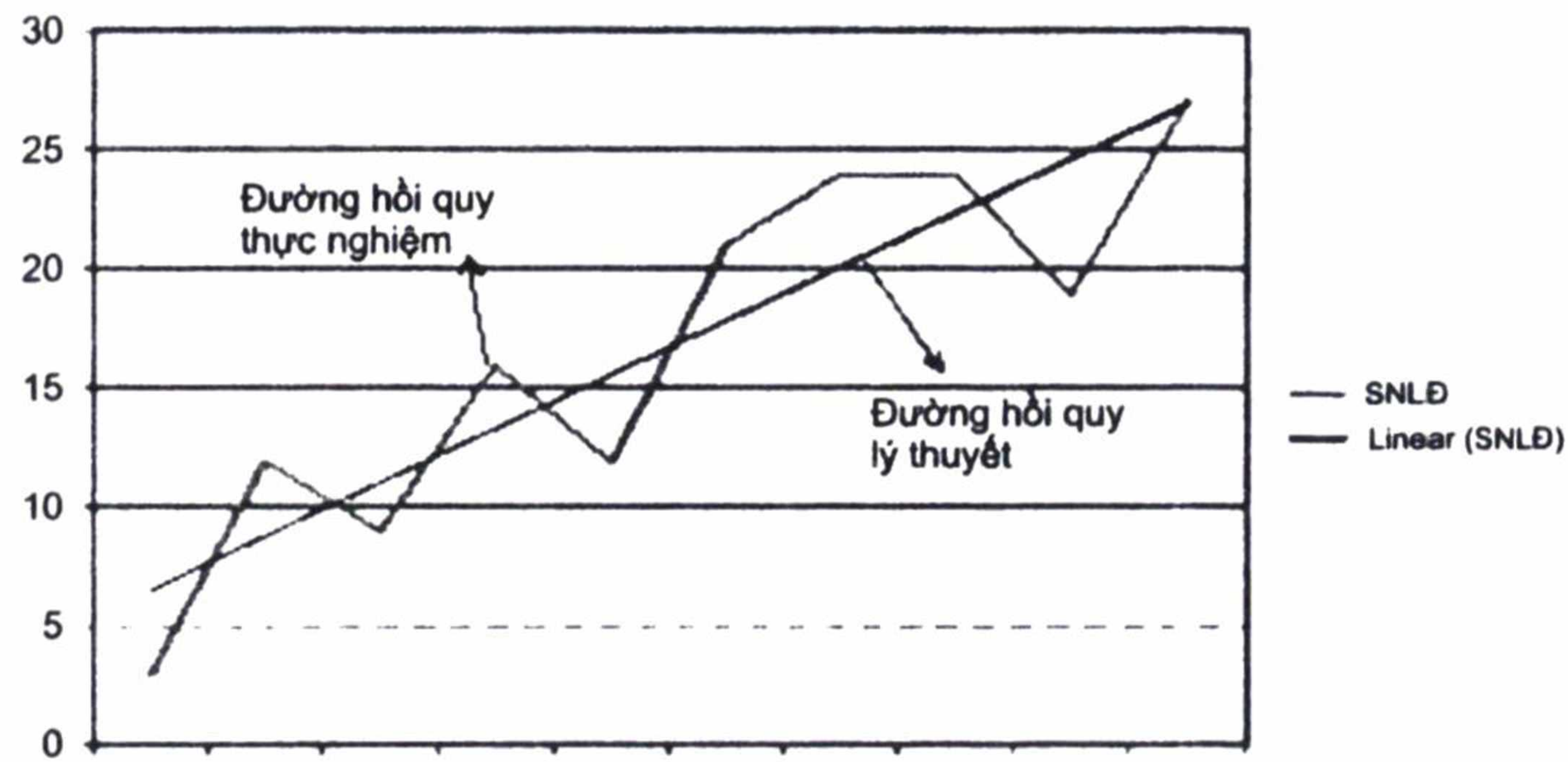
Bảng 5.1: Thâm niên công tác và tiền lương của công nhân

STT	Thâm niên công tác (năm)	Tiền lương (Tr.đ)
1	1	40
2	3	55
3	4	45
4	6	60
5	7	60
6	8	75
7	10	72
8	12	85
9	14	80
10	15	98

Qua tài liệu ta nhận thấy: Theo xu thế chung khi thâm niên công tác tăng lên, thì tiền lương của công nhân cũng có xu hướng tăng lên, nghĩa là giữa tiêu thức thâm niên công tác và tiền lương có mối liên hệ với

nhau, nhưng mỗi liên hệ này không hoàn toàn chặt chẽ, hay mỗi liên hệ tương quan có thể quan sát trên đồ thị (hình vẽ).

Đồ thị 5.1: Mỗi liên hệ giữa thâm niên công tác và tiền lương của công nhân



Đường gấp khúc biểu diễn mỗi liên hệ giữa thâm niên công tác (x) và tiền lương (y) gọi là *đường hồi quy thực nghiệm*, được hình thành từ tài liệu điều tra thực tế. Đường này chưa phản ánh rõ nét mối quan hệ giữa hai tiêu thức nhưng có xu hướng đi lên từ trái qua phải, điều này cho phép ta tìm một đường thẳng cùng hướng, gần với nó có thể thay thế cho đường hồi quy thực nghiệm nhưng biểu hiện rõ hơn mối liên hệ. Đường thẳng này được gọi là *đường hồi quy lý thuyết*, trên đồ thị vị trí của đường hồi quy lý thuyết được xác định bởi phương trình có dạng:

$$\hat{Y}_x = a + bx \tag{5.1}$$

Trong đó:

- x : Trị số của tiêu thức nguyên nhân (thâm niên công tác).
- $\hat{Y}_x$: Trị số lý thuyết của tiêu thức kết quả (tiền lương).
- a, b: Các tham số của phương trình.

Phương trình (5.1) gọi là phương trình hồi quy tuyến tính, các tham số a, b sẽ quy định vị trí của đường hồi quy lý thuyết. Các tham số này phải được xác định sao cho đường hồi quy lý thuyết có thể mô tả một cách sát nhất mỗi liên hệ tương quan. Thông thường, ta áp dụng phương

pháp bình phương nhỏ nhất nghĩa là sao cho tổng bình phương các chênh lệch giữa các trị số thực tế và trị số lý thuyết là cực tiểu:

$$\sum (y - \hat{y}_x)^2 = \min$$

Từ đó suy ra hệ phương trình chuẩn:

$$\begin{cases} na + b.\sum x = \sum y \\ a.\sum x + b.\sum x^2 = \sum x.y \end{cases} \tag{5.2}$$

Giải hệ phương trình (5.2) tính được a, b hoặc có thể biến đổi tiếp theo ta có:

$$\begin{cases} b = \frac{\overline{x.y} - \bar{x}.\bar{y}}{\sigma^2_x} \\ a = \bar{y} - b.\bar{x} \end{cases} \tag{5.3}$$

Để giải hệ phương trình (5.2) ta lập bảng tính toán:

Bảng 5.2: Các đại lượng tính các tham số của phương trình

Thâm niên công tác (x)	Tiền lương (y)	x.y	x ²	y ²
1	40	40	1	1.600
3	55	165	9	3.025
4	45	180	16	2.025
6	60	360	36	3.600
7	60	420	49	3.600
8	75	600	64	5.625
10	72	720	100	5.184
12	85	1.020	144	7.225
14	80	1.120	196	6.400
15	98	1.470	225	9.604
80	670	6.095	840	47.888

Thay số liệu trong bảng tính toán vào hệ phương trình (5.2):

$$\begin{cases} 10a + 80b = 670 \\ 80a + 840b = 6.095 \end{cases}$$

Giải hệ phương trình:

$$a = 37,6$$

$$b = 3,675$$

Tính toán a, b theo công thức (5.3) ta có:

$$\bar{x} = \frac{\sum x}{n} = \frac{80}{10} = 8 ; \bar{y} = \frac{\sum y}{n} = \frac{670}{10} = 67$$

$$\overline{xy} = \frac{\sum xy}{n} = \frac{6095}{10} = 609,5 ; \sigma_x^2 = \overline{x^2} - \bar{x}^2 = \frac{840}{10} - 8^2 = 20$$

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = \frac{609,5 - 8 \times 67}{20} = 3,675$$

$$a = \bar{y} - b \cdot \bar{x} = 67 - 3,675 \times 8 = 37,6$$

Phương trình hồi quy có dạng:

$$\hat{Y}_x = 37,6 + 3,675x$$

Trong phương trình này: Tham số $a = 37,6$ là điểm xuất phát của đường hồi quy lý thuyết, đây là tham số tự do không phụ thuộc vào x, nói lên mức độ ảnh hưởng của các nguyên nhân khác đối với tiền lương của công nhân; Tham số $b = 3,675$ gọi là *hệ số hồi quy*, quy định độ dốc của đường hồi quy lý thuyết, nói lên ảnh hưởng của tiêu thức nguyên nhân x tới tiêu thức kết quả y. Cụ thể khi thâm niên công tác tăng 1 năm thì tiền lương của công nhân tăng trung bình là 3,675 triệu đồng.

5.2.2. Hệ số tương quan

Một trong những yêu cầu quan trọng của phân tích hồi quy và tương quan là xác định cụ thể trình độ chặt chẽ của mối liên hệ. *Hệ số tương quan* là chỉ tiêu tương đối (số lần) để đánh giá trình độ chặt chẽ và chiều hướng của mối liên hệ tương quan tuyến tính.

Hệ số tương quan giúp ta xác định được cường độ của mỗi liên hệ, xem xét mức độ ảnh hưởng của tiêu thức nguyên nhân đến tiêu thức kết quả. Trong những điều kiện thời gian và không gian khác nhau, mỗi liên hệ tương quan giữa cùng một số hiện tượng cũng có thể có trình độ chặt chẽ khác nhau. Qua việc đánh giá này có thể tìm ra các nguyên nhân chủ yếu tác động đến hiện tượng nghiên cứu.

Hệ số tương quan giúp ta xác định chiều hướng của mỗi liên hệ (mỗi liên hệ thuận hay nghịch).

Hệ số tương quan cho phép kiểm định giả thuyết về sự tồn tại hay không tồn tại mỗi liên hệ tương quan tuyến tính.

Hệ số tương quan còn được sử dụng trong dự báo thống kê và tính sai số dự báo.

Trong toán học có nhiều phương pháp và nhiều công thức tính hệ số tương quan. Thống kê học vận dụng một số công thức đơn giản và thích hợp nhất. Sau đây là một số công thức tính hệ số tương quan thường dùng:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}} \quad (5.4)$$

Từ công thức (5.4) biến đổi ra một số công thức tính sau:

$$r = \frac{\overline{x.y} - \bar{x}.\bar{y}}{\sigma_x \cdot \sigma_y} \quad (5.5)$$

$$r = b \cdot \frac{\sigma_x}{\sigma_y} \quad (5.6)$$

Có thể tính hệ số tương quan từ ví dụ bằng các công thức (5.4), (5.5), (5.6).

Giả sử theo công thức (5.6):

$$\sigma_y^2 = \overline{y^2} - \bar{y}^2 = \frac{47888}{10} - 67^2 = 299,8$$

$$r = 3,675 \times \frac{\sqrt{20}}{\sqrt{299,8}} = 0,949$$

Như vậy, hệ số tương quan có thể tính bằng nhiều công thức khác nhau, tùy theo nguồn tài liệu cụ thể ta có thể sử dụng công thức tính toán phù hợp.

Hệ số tương quan có các tính chất sau:

- Hệ số tương quan nhận giá trị trong khoảng: $-1 \leq r \leq +1$
r: Mang dấu (+) mỗi liên hệ tương quan thuận.
r: Mang dấu (-) mỗi liên hệ tương quan nghịch.
- $r = \pm 1$: Giữa x và y có liên hệ hàm số.
- $r = 0$: Giữa x, y không có liên hệ tương quan tuyến tính.
- r càng gần ± 1 , mỗi liên hệ càng chặt chẽ và càng gần 0 mỗi liên hệ càng lỏng lẻo. Mức độ chặt chẽ hay lỏng lẻo của hệ số tương quan được xác định như sau:

Bảng 5.3: Mức độ chặt chẽ (lỏng lẻo) của hệ số tương quan

$ r $ Giá trị tuyệt đối của hệ số tương quan	< 0,3	0,3 - 0,7	> 0,7
Mức độ phụ thuộc	Lỏng lẻo	Trung bình	Chặt chẽ

Theo ví dụ $r = 0,949$, có thể đánh giá: Mỗi liên hệ giữa thâm niên công tác và tiền lương là mỗi liên hệ tương quan thuận và tương đối chặt chẽ.

Chú ý: Ở mục trên ta nghiên cứu mỗi liên hệ tương quan thuận giữa thâm niên công tác và tiền lương qua tài liệu điều tra trên 10 công nhân. Nhưng trong thực tế, khi nghiên cứu mối quan hệ tương quan giữa các tiêu thức cần tuân theo quy luật số lớn nghĩa là phải thu thập tài liệu ở một số đủ lớn đơn vị. Các tài liệu này sẽ được phân tổ kết hợp theo tiêu thức nguyên nhân (x) và tiêu thức kết quả (y).

Nếu gọi:

- n_x : Tần số các tổ được phân tổ theo tiêu thức x.
- n_y : Tần số các tổ được phân tổ theo tiêu thức y.
- n_{xy} : Tần số các tổ được phân tổ kết hợp theo tiêu thức x, y.
- N: Tổng tần số (số đơn vị).

$$N = \sum n_x = \sum n_y = \sum \sum n_{xy}$$

Trường hợp phân tổ có khoảng cách tổ, các giá trị x, y là trị số giữa của mỗi tổ. Ta có bảng phân tổ sau:

Bảng 5.4: Phân tổ kết hợp theo tiêu thức x và y

y \ x	x	...	n _y
	...	n _{xy}	...
	n _x	...	N

Phương trình hồi quy: $\hat{y}_x = a + bx$

Tính các tham số a, b và hệ số tương quan đều phải có tần số tương ứng.

Hệ phương trình xác định tham số a,b:

$$\begin{cases} N.a + b \sum x.n_x = \sum y.n_y \\ a. \sum x.n_x + b \sum x^2.n_x = \sum x.y.n_{xy} \end{cases} \tag{5.7}$$

Hệ số tương quan:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y}).n_{xy}}{\sqrt{\sum (x - \bar{x})^2.n_x \sum (y - \bar{y})^2.n_y}} \tag{5.8}$$

5.3. Liên hệ tương quan phi tuyến tính giữa hai tiêu thức số lượng

5.3.1. Các phương trình hồi quy phi tuyến tính

Trong nhiều trường hợp nghiên cứu ta thấy mối liên hệ tương quan giữa các tiêu thức không phải khi nào cũng được biểu hiện bằng một đường thẳng, mà có thể là các đường cong có hình dáng khác nhau, đó là mối liên hệ tương quan phi tuyến tính.

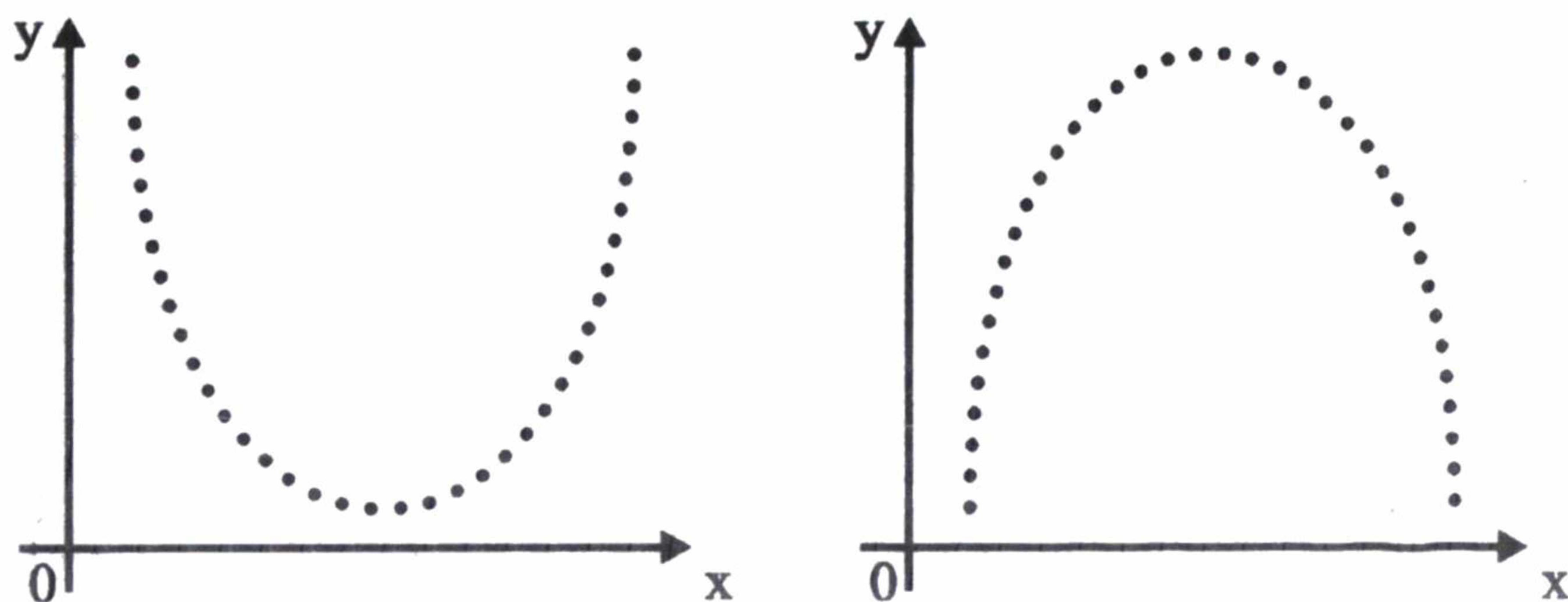
Các mối liên hệ tương quan như: Giữa tuổi nghề với năng suất lao động, giữa chi phí quảng cáo với doanh thu tiêu thụ sản phẩm..., trong

một phạm vi nghiên cứu nhất định có thể được biểu hiện là mối liên hệ tương quan tuyến tính. Nhưng thực tế, các doanh nghiệp ở các điều kiện sản xuất kinh doanh khác nhau thì mối liên hệ trên lại có thể là phi tuyến tính. Chẳng hạn: Trong một số ngành sản xuất khi tuổi nghề tăng lên thì lúc đầu làm cho năng suất lao động tăng theo, nhưng sau đó lại giảm dần. Điều đó có thể giải thích rằng: Khi tuổi nghề tăng lên đồng nghĩa với tuổi đời cao, sức lao động giảm và năng suất lao động giảm. Cũng như vậy, khi doanh nghiệp tăng chi phí cho quảng cáo làm tăng doanh thu nhưng chỉ đến mức độ nào đó thì dừng lại hoặc có tăng với mức độ chậm dần... Như vậy, nếu ta thu thập được số lớn dữ liệu và phân tích, thì mối liên hệ giữa tuổi nghề và năng suất lao động, giữa chi phí quảng cáo và doanh thu đều được biểu hiện bằng các đường cong tức là liên hệ tương quan phi tuyến tính. Ta còn có thể gặp các trường hợp khác tương tự như khi nghiên cứu mối liên hệ giữa năng suất lao động với giá thành đơn vị sản phẩm, thu nhập với nhu cầu tiêu dùng hàng hóa và dịch vụ của dân cư...

Về mặt toán học, có nhiều phương trình hồi quy để biểu hiện mối liên hệ tương quan phi tuyến tính. Sau đây là một số dạng phương trình tiêu biểu:

- *Phương trình hồi quy pa-ra-bôn:*

Trường hợp khi tiêu thức nguyên nhân tăng (giảm) với lượng đều nhau, thì tiêu thức kết quả biến động không đều và đến một mức độ nào đó (cực tiểu hay cực đại) lại đảo chiều. Ví dụ như mối liên hệ giữa tuổi nghề và năng suất lao động, thu nhập và nhu cầu tiêu dùng... Nếu thăm dò bằng đồ thị trực hoành là tiêu thức nguyên nhân (x), trục tung là tiêu thức kết quả (y), các điểm trên đồ thị được phân bố theo một trong hai dạng sau đây có thể xây dựng mô hình hồi quy pa-ra-bôn:



Phương trình pa-ra-bôn:

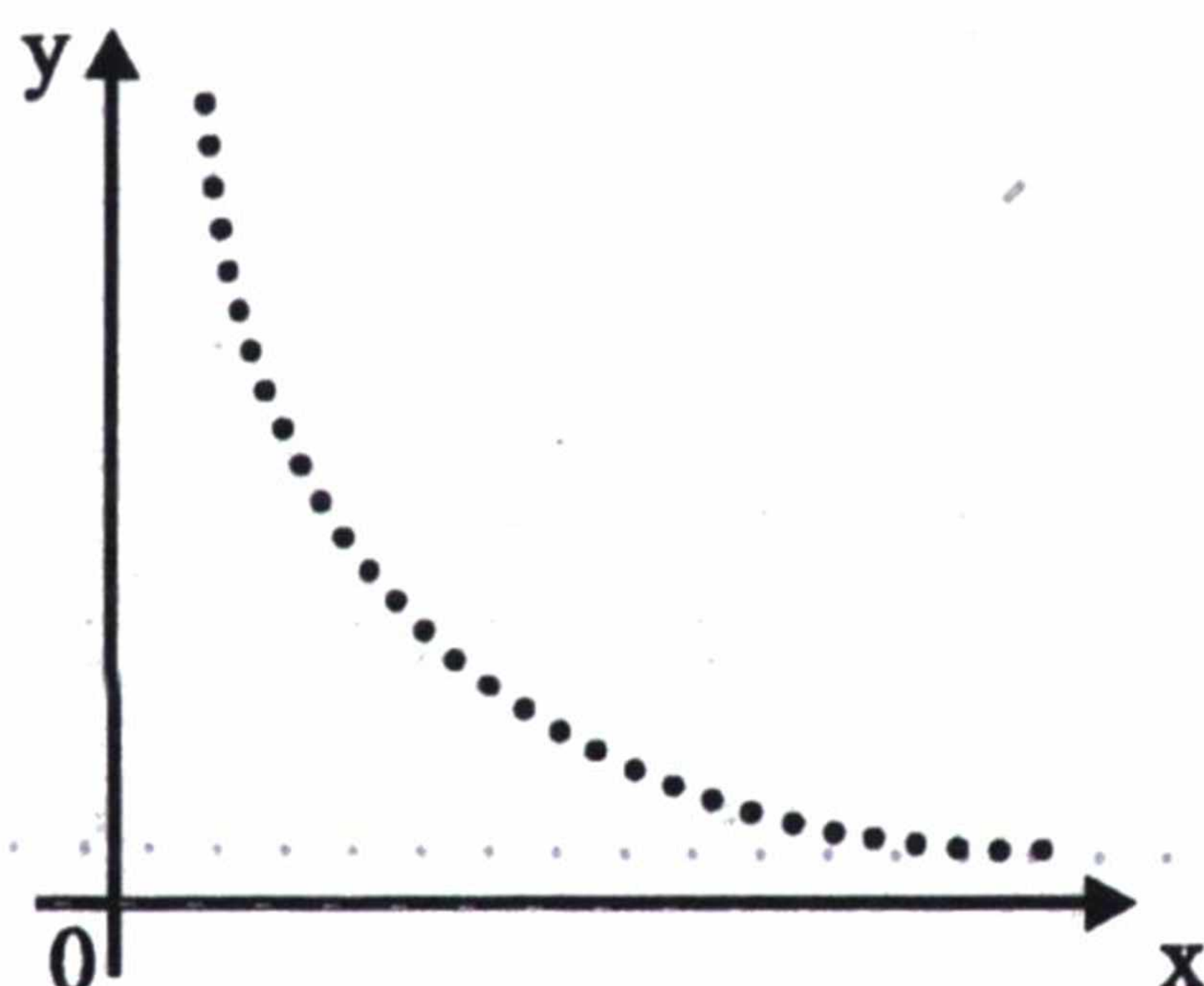
$$\hat{y}_x = a + bx + cx^2 \quad (5.9)$$

Áp dụng phương pháp bình phương nhỏ nhất sẽ có hệ phương trình sau đây để tìm giá trị các hệ số a, b, c:

$$\begin{cases} \sum y = na + b \sum x + c \sum x^2 \\ \sum xy = a \sum x + b \sum x^2 + c \sum x^3 \\ \sum x^2 y = a \sum x^2 + b \sum x^3 + c \sum x^4 \end{cases} \quad (5.10)$$

- *Phương trình hồi quy hy-pe-bôn:*

Trường hợp khi tiêu thức nguyên nhân tăng mà trị số của tiêu thức kết quả giảm với tốc độ không đều, lúc đầu nhanh sau chậm dần. Ví dụ như mối liên hệ giữa sản lượng sản xuất với giá thành đơn vị sản phẩm. Nếu thăm dò bằng đồ thị, các điểm trên đồ thị được phân bố theo dạng sau đây có thể xây dựng mô hình hồi quy hy-pe-bôn:



Phương trình hy-pe-bôn:

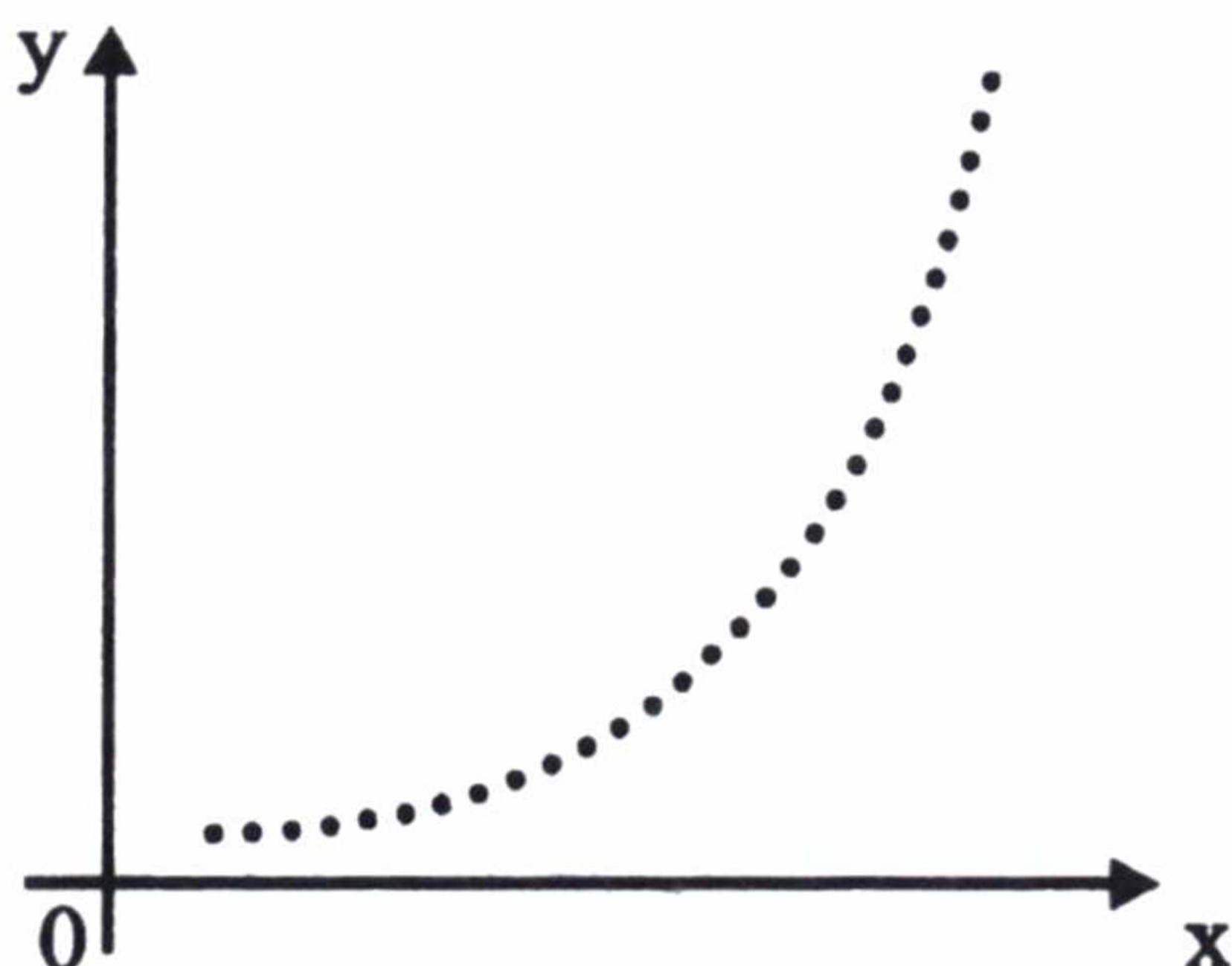
$$\hat{y}_x = a + \frac{b}{x} \quad (5.11)$$

Áp dụng phương pháp bình phương nhỏ nhất sẽ có hệ phương trình sau đây để tìm giá trị các hệ số a, b:

$$\begin{cases} \sum y = na + b \sum \frac{1}{x} \\ \sum y \frac{1}{x} = a \sum \frac{1}{x} + b \sum \frac{1}{x^2} \end{cases} \quad (5.12)$$

- *Phương trình hàm số mũ:*

Trường hợp khi tiêu thức nguyên nhân tăng (giảm) làm cho tiêu thức kết quả thay đổi gần như một cấp số nhân. Nếu thăm dò bằng đồ thị các điểm trên đồ thị được phân bố theo dạng sau đây có thể xây dựng mô hình hàm mũ:



Phương trình hàm số mũ:

$$\hat{y}_x = ab^x \quad (5.13)$$

Áp dụng phương pháp bình phương nhỏ nhất sẽ có hệ phương trình sau đây để tìm giá trị các hệ số a, b :

$$\begin{cases} \sum \ln y = n \ln a + \ln b \sum x \\ \sum x \ln y = \ln a \sum x + \ln b \sum x^2 \end{cases} \quad (5.14)$$

Giải hệ phương trình trên sẽ được $\ln a, \ln b$. Tra đối $\ln$ sẽ tìm được giá trị của a, b .

Ngoài ra, các mối liên hệ còn có thể biểu hiện dưới dạng hàm bậc 3, hàm lũy thừa, hàm logarit... Việc vận dụng dạng phương trình nào phải căn cứ vào việc phân tích tính chất của mối liên hệ giữa các tiêu thức.

5.3.2. Tỷ số tương quan (Ký hiệu η : êta)

Tỷ số tương quan được sử dụng để đánh giá mức độ chặt chẽ mối liên hệ tương quan phi tuyến tính giữa hai tiêu thức số lượng và được tính theo công thức sau đây:

$$\eta = \sqrt{1 - \frac{\sum (y - \hat{y}_x)^2}{\sum (y - \bar{y})^2}} \quad (5.15)$$

Trong đó : y : Các giá trị thực tế.

$\hat{y}_x$: Các giá trị lý thuyết của y theo x .

$\bar{y}$: Giá trị trung bình của y .

Tính chất: η nằm trong khoảng $[0;1]$ tức là: $0 \leq \eta \leq 1$

Cụ thể:

- nếu $\eta = 1$: Giữa x và y có mối liên hệ hàm số.
- Nếu $\eta = 0$: Giữa x và y không có mối liên hệ.
- Nếu η càng gần 1 thì mối liên hệ càng chặt chẽ và càng gần 0 thì mối liên hệ càng lỏng lẻo.

Tỷ số tương quan có thể đánh giá được trình độ chặt chẽ của mối liên hệ, nhưng không đánh giá được chiều hướng của mối liên hệ.

Tỷ số tương quan còn có thể được sử dụng trong trường hợp đánh giá mức độ chặt chẽ của các mối liên hệ tương quan tuyến tính. Khi giữa các tiêu thức x và y có liên hệ tương quan tuyến tính thì việc tính r và η đều cho kết quả giống nhau. Còn nếu x và y có liên hệ tương quan phi tuyến tính thì các kết quả tính toán giữa hai chỉ tiêu này sẽ khác nhau.

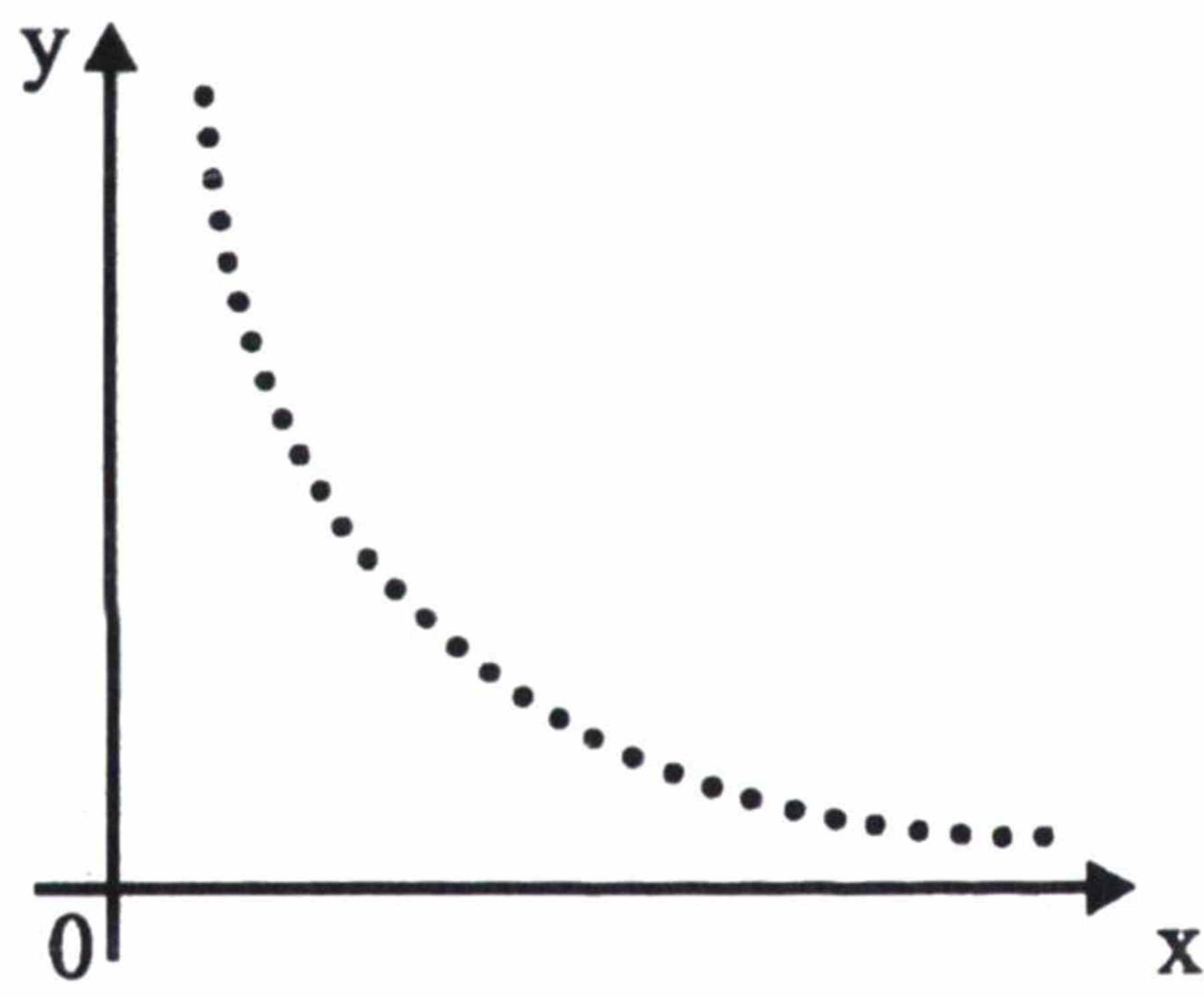
Sau đây là ví dụ phân tích tương quan phi tuyến tính: Có tài liệu sản lượng sản xuất (nghìn sản phẩm) và giá thành đơn vị sản phẩm (nghìn đồng/sản phẩm) của 10 doanh nghiệp cùng sản xuất một loại sản phẩm như sau:

Bảng 5.5: Sản lượng sản xuất và giá thành đơn vị sản phẩm

Sản lượng (1000.sp)	Giá thành (1000.đ/sp)	Sản lượng (1000.sp)	Giá thành (1000.đ/sp)
10	15,60	35	15,15
15	15,40	40	15,14
20	15,27	50	15,12
25	15,24	60	15,10
30	15,20	80	15,05

Qua quan sát tài liệu cho thấy: Sản lượng tăng lên thì giá thành đơn vị sản phẩm giảm xuống, nhưng tốc độ giảm không đều, mà giảm chậm dần. Đó là do trong giá thành sản phẩm bao gồm hai bộ phận chi phí: Chi phí khả biến và chi phí bất biến. Chi phí khả biến phụ thuộc vào sản lượng sản xuất của doanh nghiệp: nếu ký hiệu sản lượng là (x), chi phí khả biến cho một đơn vị sản phẩm là (a), thì tổng chi phí khả biến sản xuất sản phẩm là (a.x). Phần chi phí bất biến không phụ thuộc vào sản lượng, ký hiệu là (b). Vậy, tổng chi phí sản xuất sản phẩm (hay tổng giá thành) là: (a.x + b) và giá thành đơn vị sản phẩm sẽ là: $(a + \frac{b}{x})$.

Nếu thăm dò trên đồ thị, tiêu thức nguyên nhân là sản lượng (x), tiêu thức kết quả là giá thành đơn vị sản phẩm (y) và có đồ thị sau đây:



Từ đồ thị trên, có thể xây dựng mô hình hy-pe-bôn để biểu hiện mối liên hệ giữa sản lượng và giá thành đơn vị sản phẩm: $\hat{y}_x = a + \frac{b}{x}$.

Căn cứ vào hệ phương trình của mô hình hy-pe-bôn, ta có bảng tính toán sau đây:

Bảng 5.6: Các đại lượng để tính các tham số của phương trình

x	y	$\frac{1}{x}$	$\frac{1}{x^2}$	$y\frac{1}{x}$
10	15,60	0,100	0,010	1,560
15	15,40	0,067	0,004	1,032
20	15,27	0,050	0,003	0,764
25	15,24	0,040	0,002	0,610
30	15,20	0,033	0,001	0,502
35	15,15	0,029	0,001	0,439
40	15,14	0,025	0,001	0,379
50	15,12	0,020	0,001	0,302
60	15,10	0,017	0,001	0,257
80	15,05	0,013	0,001	0,196
	$\Sigma y = 152,27$	$\Sigma \frac{1}{x} = 0,394$	$\Sigma \frac{1}{x^2} = 0,025$	$\Sigma y\frac{1}{x} = 6,041$

Thay số liệu vào hệ phương trình:

$$\begin{cases} 152,27 = 10a + 0,394b \\ 6,041 = 0,394a + 0,025b \end{cases}$$

Giải hệ phương trình ta có: $a = 15,054$; $b = 4,385$

Phương trình hồi quy:

$$\hat{y} = 15,054 + \frac{4,385}{x}$$

Bằng cách lập bảng tương tự để tính tỷ số tương quan:

Bảng 5.7: Các đại lượng để tính tỷ số tương quan

x	y	$\hat{y}_x$	$(y - \hat{y}_x)^2$	$(y - \bar{y})^2$
10	15,60	15,49	0,0121	0,1391
15	15,40	15,35	0,0025	0,0299
20	15,27	15,27	0,0001	0,0018
25	15,24	15,23	0,0001	0,0001
30	15,20	15,20	0,0000	0,0007
35	15,15	15,18	0,0009	0,0059
40	15,14	15,16	0,0004	0,0076
50	15,12	15,14	0,0004	0,0114
60	15,10	15,13	0,0009	0,0161
80	15,05	15,11	0,0036	0,0313
Σ		152,26	0,0209	0,2442

Từ số liệu bảng trên, áp dụng công thức (5.15) tính được tỷ số tương quan:

$$\eta = \sqrt{1 - \frac{0,0209}{0,2442}} = 0,956$$

Như vậy, mối liên hệ giữa sản lượng và giá thành đơn vị sản phẩm khá chặt chẽ.

5.4. Liên hệ tương quan tuyến tính giữa nhiều tiêu thức số lượng

Trong các phần trên ta đã nghiên cứu mối liên hệ giữa hai tiêu thức: Một tiêu thức nguyên nhân, một tiêu thức kết quả. Song, trên thực tế mối liên hệ thường có nhiều tiêu thức nguyên nhân đồng thời tác động đến một kết quả. Ví dụ như: Nhu cầu tiêu dùng có thể do ảnh hưởng của một số nhân tố như: Thu nhập, giá cả hàng hóa, điều kiện tự nhiên, phong tục tập quán, sở thích, thói quen của người tiêu dùng, các chính sách kinh tế - xã hội của chính phủ... Vận dụng phương pháp hồi quy tương quan để

Giả sử có hai tiêu thức nguyên nhân x_1, x_2 thì phương trình hồi quy:

$$\hat{y}_x = a_0 + a_1x_1 + a_2x_2 \quad (5.18)$$

Và hệ phương trình xác định a_0, a_1, a_2

$$\begin{cases} a_0n + a_1 \cdot \sum x_1 + a_2 \sum x_2 = \sum y \\ a_0 \cdot \sum x_1 + a_1 \sum x_1^2 + a_2 \sum x_1x_2 = \sum x_1y \\ a_0 \cdot \sum x_2 + a_1 \sum x_1x_2 + a_2 \sum x_2^2 = \sum x_2y \end{cases} \quad (5.19)$$

Các tham số a_0, a_1, a_2 cũng có thể được tính từ công thức:

$$\begin{aligned} a_1 &= \frac{\sigma_y}{\sigma_{x_1}} \times \frac{r_{yx_1} - r_{yx_2} \times r_{x_1x_2}}{1 - r_{x_1x_2}^2} \\ a_2 &= \frac{\sigma_y}{\sigma_{x_2}} \times \frac{r_{yx_2} - r_{yx_1} \times r_{x_1x_2}}{1 - r_{x_1x_2}^2} \\ a_0 &= \bar{y} - a_1 \bar{x}_1 - a_2 \bar{x}_2 \end{aligned} \quad (5.20)$$

Trong đó: $r_{yx_1}, r_{yx_2}, r_{x_1x_2}$ là các hệ số tương quan tuyến tính giữa các cặp tiêu thức yx_1, yx_2 và x_1x_2 :

$$\begin{aligned} r_{yx_1} &= \frac{\overline{yx_1} - \bar{y} \cdot \bar{x}_1}{\sigma_y \sigma_{x_1}} \\ r_{yx_2} &= \frac{\overline{yx_2} - \bar{y} \cdot \bar{x}_2}{\sigma_y \sigma_{x_2}} \\ r_{x_1x_2} &= \frac{\overline{x_1x_2} - \bar{x}_1 \cdot \bar{x}_2}{\sigma_{x_1} \sigma_{x_2}} \end{aligned} \quad (5.21)$$

5.4.2. Hệ số tương quan

Để đánh giá trình độ chặt chẽ của mối liên hệ giữa nhiều tiêu thức ta tính hệ số tương quan bội và hệ số tương quan riêng.

Hệ số tương quan bội được dùng để đánh giá trình độ chặt chẽ của mối liên hệ giữa tất cả các tiêu thức nguyên nhân nghiên cứu đến tiêu thức kết quả.

Hệ số tương quan riêng: Được dùng để đánh giá trình độ chặt chẽ của mỗi liên hệ riêng từng tiêu thức nguyên nhân đến tiêu thức kết quả với điều kiện đã loại trừ ảnh hưởng của các tiêu thức nguyên nhân khác.

Hệ số tương quan bội giữa hai tiêu thức nguyên nhân x_1, x_2 với tiêu thức kết quả y được tính theo công thức:

$$R_{yx_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}} \tag{5.22}$$

Hệ số tương quan bội nhận giá trị từ $0 \leq R \leq 1$

$R = 0$: Không có liên hệ tuyến tính.

$R = 1$: Mỗi liên hệ hàm số.

R càng gần 1 thì mỗi liên hệ càng chặt chẽ.

Để phân tích sâu sắc hơn, có thể đánh giá trình độ chặt chẽ của mỗi liên hệ riêng từng tiêu thức nguyên nhân x_1, x_2 đến tiêu thức kết quả y , với điều kiện loại trừ ảnh hưởng của nguyên nhân khác. Đó là các hệ số tương quan riêng:

Hệ số tương quan riêng giữa x_1 với y (loại trừ ảnh hưởng của x_2).

$$R_{yx_1(x_2)} = \frac{r_{yx_1} - r_{yx_2}r_{x_1x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1x_2}^2)}} \tag{5.23}$$

Hệ số tương quan riêng giữa x_2 với y (loại trừ ảnh hưởng của x_1).

$$R_{yx_2(x_1)} = \frac{r_{yx_2} - r_{yx_1}r_{x_1x_2}}{\sqrt{(1 - r_{yx_1}^2)(1 - r_{x_1x_2}^2)}} \tag{5.24}$$

Sau đây là ví dụ phân tích mỗi liên hệ tương quan giữa ba tiêu thức:

Giả sử có tài liệu điều tra thống kê về tuổi đời, tuổi nghề và năng suất lao động của 5 công nhân trong một doanh nghiệp như sau:

Bảng 5.8: Số liệu về tuổi nghề, tuổi đời, năng suất lao động của công nhân

Công nhân	Tuổi nghề (năm) x_1	Tuổi đời (năm) x_2	Năng suất lao động (sản phẩm) y
A	1	22	4
B	3	21	5
C	4	24	6
D	8	26	7
E	10	35	9
Tổng	26	128	31
Số trung bình	5,2	25,6	6,2

Phương trình hồi quy biểu hiện mối liên hệ giữa 3 tiêu thức tuổi nghề (x_1), tuổi đời (x_2) và năng suất lao động (y). Có dạng như sau:

$$\hat{y}_x = a_0 + a_1x + a_1x_2$$

Để tính các tham số a_0, a_1, a_2 ta lập bảng tính toán sau:

Bảng 5.9: Các đại lượng để tính các tham số của phương trình

Công nhân	x_1y	x_2	x_1x_2	x_1^2	x_2^2	y^2
A	4	88	22	1	484	16
B	15	105	63	9	441	25
C	24	144	96	16	576	36
D	56	182	208	64	676	49
E	90	315	350	100	1225	81
Cộng	189	834	739	190	3402	207
Số trung bình	37,8	166,8	147,8	38	680,4	41,4

Theo tài liệu ở bảng trên, tính được:

$$\sigma_{x_1} = \sqrt{x_1^2 - \bar{x}_1^2} = \sqrt{38 - 5,2^2} = 3,31$$

$$\sigma_{x_2} = \sqrt{x_2^2 - \bar{x}_2^2} = \sqrt{680,4 - 25,6^2} = 5,00$$

$$\sigma_y = \sqrt{y^2 - \bar{y}^2} = \sqrt{41,4 - 6,2^2} = 1,72$$

$$r_{yx_1} = \frac{\overline{yx_1} - \bar{y} \cdot \bar{x}_1}{\sigma_y \sigma_{x_1}} = \frac{37,8 - 6,2 \times 5,2}{1,72 \times 3,31} = 0,976$$

$$r_{yx_2} = \frac{\overline{yx_2} - \bar{y} \cdot \bar{x}_2}{\sigma_y \sigma_{x_2}} = \frac{166,8 - 6,2 \times 25,6}{1,72 \times 5} = 0,94$$

$$r_{x_1 x_2} = \frac{\overline{x_1 x_2} - \bar{x}_1 \cdot \bar{x}_2}{\sigma_{x_1} \sigma_{x_2}} = \frac{147,8 - 5,2 \times 25,6}{3,31 \times 5,0} = 0,887$$

Từ kết quả trên tính ra các tham số của phương trình hồi quy theo công thức (5.20):

$$a_1 = \frac{1,72}{3,31} \times \frac{0,976 - 0,94 \times 0,887}{1 - 0,887^2} = 0,35$$

$$a_2 = \frac{1,72}{5,00} \times \frac{0,94 - 0,976 \times 0,887}{1 - 0,887^2} = 0,12$$

$$a_0 = 6,2 - 0,35 \times 5,2 - 0,12 \times 25,6 = 1,31$$

Ta có phương trình hồi quy:

$$y_x = 1,31 + 0,35x_1 + 0,12x_2$$

Hệ số tương quan bội tính theo công thức (5.22):

$$R_{y_{x_1 x_2}} = \sqrt{\frac{0,976^2 + 0,94^2 - 2 \times 0,976 \times 0,94 \times 0,887}{1 - 0,887^2}} = 0,99$$

Các hệ số tương quan riêng được tính theo công thức (5.23), (5.24):

- Hệ số tương quan riêng giữa tuổi nghề (x_1) và năng suất lao động (y), loại trừ ảnh hưởng của tuổi đời (x_2):

$$R_{y_{x_1(x_2)}} = \frac{0,976 - 0,94 \times 0,887}{\sqrt{(1 - 0,94^2)(1 - 0,887^2)}} = 0,90$$

- Hệ số tương quan riêng giữa tuổi đời (x_2) và năng suất lao động (y), loại trừ ảnh hưởng của tuổi nghề (x_1):

$$R_{y_{x_2}(x_1)} = \frac{0,94 - 0,976 \times 0,887}{\sqrt{(1 - 0,976^2)(1 - 0,887^2)}} = 0,74$$

Theo các kết quả tính toán ở trên mỗi liên hệ giữa tuổi nghề, tuổi đời với năng suất lao động của công nhân trong doanh nghiệp trên là mỗi liên hệ tương quan thuận và rất chặt chẽ ($R_{yx_1x_2} = 0,99$). Trong đó, mỗi liên hệ giữa tuổi nghề với năng suất lao động chặt chẽ hơn mỗi liên hệ giữa tuổi đời với năng suất lao động ($R_{y_{x_1}(x_2)} > R_{y_{x_2}(x_1)}$).

5.5. Hệ số co giãn

Phương pháp hồi quy và tương quan giúp ta nghiên cứu mỗi liên hệ giữa các tiêu thức, ngoài ra còn có thể tính toán *hệ số co giãn* để đo mức độ phản ứng của tiêu thức kết quả (y) đối với sự biến thiên của tiêu thức nguyên nhân (x).

Giả sử phương trình hồi quy $\hat{y}_x = f(x)$; số gia của tiêu thức nguyên nhân x là Δx , số gia của tiêu thức kết quả y là $\Delta y = f(x + \Delta x) - f(x)$, ta có hệ số co giãn:

$$E = \frac{\Delta y}{y} \div \frac{\Delta x}{x} \quad (5.25)$$

Trong trường hợp x, y có mối quan hệ tương quan tuyến tính thì:

$$E = b \frac{x}{y} \quad (5.26)$$

b: Hệ số hồi quy

Thông thường, người ta thay giá trị x, y bằng số trung bình của chúng, khi đó:

$$E = b \frac{\bar{x}}{\bar{y}} \quad (5.27)$$

Hệ số co giãn phản ánh khi tiêu thức nguyên nhân x thay đổi 1% thì làm cho tiêu thức kết quả y thay đổi bao nhiêu %.

Trở lại ví dụ trong mục 5.2.

$$E = 3,675 \times \frac{8}{67} = 0,439$$

Nghĩa là khi thâm niên công tác tăng 1% thì tiền lương tăng bình quân là 0,439%.

Tính chất của hệ số co giãn:

- E dương thì giữa x,y đồng biến. E âm thì x, y nghịch biến.
- $E = \pm 1$: Sự biến thiên của x, y trùng nhau.
- $E = 0$ thì y không biến đổi (hằng số).
- $|E| > 1$: Sự biến thiên của x nhanh hơn sự biến thiên của y.
- $|E| < 1$: Sự biến thiên của y nhanh hơn sự biến thiên của x.

TÓM TẮT CHƯƠNG 5

Các hiện tượng kinh tế - xã hội luôn tồn tại trong mối liên hệ phụ thuộc và ràng buộc lẫn nhau. Tùy theo mức độ phụ thuộc của mỗi liên hệ, có thể chia thành hai loại: Liên hệ hàm số và liên hệ tương quan.

Mối liên hệ giữa các hiện tượng kinh tế - xã hội thường là mối liên hệ tương quan. Thống kê sử dụng phương pháp hồi quy và phương pháp tương quan để nghiên cứu mối liên hệ đó (có thể gọi ngắn gọn là phương pháp tương quan).

Phương pháp hồi quy và tương quan nhằm giải quyết hai nhiệm vụ cơ bản: Xác định mô hình hồi quy biểu hiện mối liên hệ và đánh giá trình độ chặt chẽ của mối liên hệ bằng các hệ số tương quan hay tỷ số tương quan.

Trường hợp giản đơn là nghiên cứu mối liên hệ giữa hai tiêu thức số lượng. Có thể dựa trên cơ sở phân tích lý luận, kết hợp các phương pháp thống kê như phương pháp đồ thị, phân tổ,... để xác định dạng mô

hình hồi quy là tuyến tính (phương trình bậc nhất) hay phi tuyến tính (phương trình bậc 2, 3, hàm số mũ, ...). Dạng tổng quát của phương trình hồi quy biểu diễn mối liên hệ giữa tiêu thức nguyên nhân (x) và tiêu thức kết quả (y) là: $y_x = f(x)$. Các hệ số của phương trình được xác định bằng phương pháp bình phương nhỏ nhất, từ đó dẫn đến hệ phương trình chuẩn, giải hệ phương trình sẽ có kết quả. Các hệ số đó sẽ cho thấy mức độ ảnh hưởng của tiêu thức nguyên nhân (x) đến tiêu thức kết quả (y). Hệ số tương quan (trường hợp mối liên hệ tương quan tuyến tính), tỷ số tương quan (trường hợp mối liên hệ tương quan phi tuyến tính) được sử dụng để đánh giá trình độ chặt chẽ của mối liên hệ tương quan giữa hai tiêu thức số lượng.

Trường hợp phức tạp là nghiên cứu mối liên hệ tương quan giữa nhiều tiêu thức (nhiều tiêu thức nguyên nhân ảnh hưởng tới một tiêu thức kết quả). Dạng tổng quát của phương trình hồi quy là: $y_x = f(x_1, x_2, \dots, x_n)$. Các hệ số của phương trình tính được sẽ đánh giá vai trò ảnh hưởng của mỗi tiêu thức nguyên nhân đến tiêu thức kết quả. Hệ số tương quan bội đánh giá trình độ chặt chẽ của mối liên hệ giữa tất cả các tiêu thức nguyên nhân đến tiêu thức kết quả. Hệ số tương quan riêng đánh giá trình độ chặt chẽ của mối liên hệ riêng từng tiêu thức nguyên nhân đến tiêu thức kết quả với điều kiện đã loại trừ ảnh hưởng của các tiêu thức nguyên nhân khác.

Phương pháp hồi quy và tương quan giúp ta nghiên cứu mối liên hệ giữa các tiêu thức. Ngoài ra, có thể tính toán hệ số co giãn để đo mức độ phản ứng của tiêu thức kết quả (y) đối với sự biến thiên của tiêu thức nguyên nhân (x).

CÂU HỎI ÔN TẬP CHƯƠNG 5

1. Ý nghĩa nghiên cứu mối liên hệ giữa các hiện tượng kinh tế - xã hội. Phân biệt các mối liên hệ hàm số và liên hệ tương quan?
2. Nhiệm vụ của phương pháp phân tích hồi quy và tương quan?
3. Ý nghĩa của các hệ số trong phương trình hồi quy tuyến tính giữa hai tiêu thức số lượng?

- 4. Ý nghĩa, phương pháp xác định, tính chất của hệ số tương quan?
- 5. Các dạng mô hình hồi quy phi tuyến tính tiêu biểu? Ý nghĩa, phương pháp tính, tính chất của tỷ số tương quan?
- 6. Ý nghĩa nghiên cứu mối liên hệ tương quan giữa nhiều tiêu thức số lượng, ý nghĩa hệ số tương quan bội và hệ số tương quan riêng?
- 7. Phương pháp tính, ý nghĩa của hệ số co giãn?

BÀI TẬP CHƯƠNG 5

Bài 1:

Có tài liệu điều tra ở một tổ sản xuất của một doanh nghiệp như sau:

Tên công nhân	Tuổi nghề (năm)	Năng suất lao động (tấn)
A	1	3
B	3	12
C	4	9
D	5	16
E	7	12
G	8	21
H	9	24
I	10	24
K	11	19
L	12	27

Yêu cầu:

- a) Trình bày mối liên hệ trên đồ thị. Rút ra nhận xét?
- b) Xác định phương trình đường thẳng biểu hiện mối quan hệ tương quan giữa tuổi nghề và năng suất lao động của công nhân tổ sản xuất trên?
- c) Đánh giá trình độ chặt chẽ của mối quan hệ đó?

Bài 2:

Có tài liệu tại một công ty thương mại như sau:

Cửa hàng	Năng suất lao động (trđ)	Tiền lương trung bình (trđ.)
Số 1	500	10
Số 2	520	12
Số 3	550	16
Số 4	560	15
Số 5	600	18

Yêu cầu:

- a) Xác định phương trình tuyến tính biểu hiện mối liên hệ tương quan giữa năng suất lao động và tiền lương trung bình của nhân viên công ty trên?
- b) Tính toán, nêu ý nghĩa của hệ số tương quan?

Bài 3:

Tại một công ty thương mại X có số liệu sau:

Cửa hàng	Doanh thu bán hàng (Tr.đ)	Chi phí kinh doanh (Tr.đ)
1	4,50	35
2	5,00	40
3	4,20	44
4	6,80	51
5	7,50	62
6	8,00	68

Yêu cầu:

- a) Xác định phương trình đường thẳng phản ánh mối liên hệ giữa doanh thu bán hàng và chi phí kinh doanh?
- b) Tính hệ số tương quan để đánh giá trình độ chặt chẽ của mối liên hệ trên?

Bài 4:

Có tài liệu thống kê về năng suất lao động trung bình và giá thành đơn vị sản phẩm của 5 phân xưởng trong doanh nghiệp như sau:

Phân xưởng	Năng suất lao động trung bình một công nhân (Tấn)	Giá thành đơn vị sản phẩm (Trđ/tấn)
A	20	18
B	22	17
C	25	15,8
D	28	16
E	30	14,2

Yêu cầu:

- a) Xác định phương trình hồi quy biểu hiện mối liên hệ giữa năng suất lao động và giá thành đơn vị sản phẩm của doanh nghiệp trên?
- b) Đánh giá trình độ chặt chẽ của mối liên hệ?

Bài 5:

Có tài liệu thống kê về 8 doanh nghiệp thương mại như sau:

Doanh nghiệp	Số lao động (người)	Vốn kinh doanh (tỷ đồng)	Doanh thu (tỷ đồng)
A	20	1,5	5
B	26	2,8	8
C	35	4,6	15
D	50	4,0	12
E	62	3,5	18
G	70	8,2	25
H	74	5,6	24
I	79	13,8	29

Yêu cầu:

- a) Xác định mô hình hồi quy phản ánh mối liên hệ giữa các tiêu thức số lao động, vốn kinh doanh và doanh thu của doanh nghiệp?
- b) Đánh giá vai trò của từng tiêu thức nguyên nhân tới tiêu thức kết quả?
- c) Đánh giá trình độ chặt chẽ của mối liên hệ?

Chương 6

DÃY SỐ THỜI GIAN

Để phân tích sự biến động hiện tượng theo thời gian cần có các dữ liệu thống kê về hiện tượng và được sắp xếp theo thứ tự thời gian gọi là dãy số thời gian. Trong chương này sẽ trình bày phương pháp phân tích dãy số thời gian nhằm biểu hiện đặc điểm, xu hướng biến động và dự báo mức độ tương lai của hiện tượng. Nội dung chương bao gồm: Khái niệm, ý nghĩa dãy số thời gian, các chỉ tiêu phân tích dãy số thời gian, các phương pháp biểu hiện xu hướng phát triển cơ bản của hiện tượng và một số phương pháp dự báo thống kê.

6.1. Khái niệm, ý nghĩa của dãy số thời gian

Các hiện tượng kinh tế - xã hội không ngừng biến động theo thời gian. Để nghiên cứu sự biến động đó, thống kê thường sử dụng phương pháp dãy số thời gian.

Dãy số thời gian là dãy các trị số của chỉ tiêu thống kê được sắp xếp theo thứ tự thời gian.

Ví dụ: Có tài liệu về doanh thu của một doanh nghiệp qua các năm như sau:

Bảng 6.1: Dãy số thời gian về doanh thu

Năm	2011	2012	2013	2014	2015
Doanh thu (tỷ.đ)	200	240	276	300	320

Dãy số trên phản ánh sự biến động chỉ tiêu doanh thu của doanh nghiệp từ năm 2011 đến năm 2015.

Một dãy số thời gian gồm hai thành phần: Thời gian và chỉ tiêu về hiện tượng nghiên cứu.

Thời gian: Tùy theo yêu cầu và mục đích nghiên cứu thời gian có thể là giờ, ngày, tháng, quý, năm... Độ dài thời gian giữa hai mức độ liên nhau được gọi là khoảng cách thời gian.

Chỉ tiêu về hiện tượng nghiên cứu thường được biểu hiện bằng các trị số cụ thể, được gọi là các mức độ của dãy số. Các mức độ này có thể là các số tuyệt đối, số tương đối hoặc số trung bình. Khi thời gian thay đổi thì các mức độ của dãy số cũng thay đổi.

Dãy số thời gian cho phép thống kê nghiên cứu đặc điểm và xu hướng về sự biến động của hiện tượng theo thời gian. Trên cơ sở đó dự báo sự phát triển của hiện tượng trong tương lai.

Căn cứ vào đặc điểm tồn tại của hiện tượng theo thời gian, có thể chia dãy số thời gian làm hai loại: *Dãy số thời kỳ* và *dãy số thời điểm*.

Dãy số thời kỳ là dãy số biểu hiện quy mô (khối lượng) của hiện tượng trong một độ dài thời gian nhất định. Các mức độ của dãy số thời kỳ là các số tuyệt đối thời kỳ nên phụ thuộc vào khoảng cách thời gian. Khoảng cách thời gian trong dãy số càng dài thì trị số của chỉ tiêu càng lớn, do đó có thể cộng các trị số của chỉ tiêu để phản ánh quy mô hiện tượng nghiên cứu trong thời kỳ dài hơn.

Dãy số thời gian về doanh thu của doanh nghiệp trên từ năm 2011 đến năm 2015, phản ánh kết quả kinh doanh của doanh nghiệp trong từng năm, là dãy số thời kỳ.

Dãy số thời điểm: Là dãy số biểu hiện quy mô (khối lượng) của hiện tượng nghiên cứu tại những thời điểm nhất định. Các mức độ của dãy số thời điểm là các số tuyệt đối thời điểm. Trong dãy số thời điểm các trị số của chỉ tiêu không phụ thuộc vào khoảng cách thời gian, không được cộng các trị số này lại với nhau.

Dãy số thời điểm có khoảng cách thời gian giữa hai thời điểm bằng nhau hoặc có thể không bằng nhau.

Ví dụ: Có tài liệu về giá trị hàng hóa tồn kho của doanh nghiệp X vào ngày đầu các tháng như sau:

Bảng 6.2: Dãy số thời gian về giá trị hàng hóa tồn kho

Ngày	1/1	1/2	1/3	1/4
Giá trị hàng hóa tồn kho (tỷ.đ)	20	22	28	36

Ở ví dụ trên các khoảng cách thời gian đều bằng nhau là một tháng (với giả sử các tháng đều có số ngày là như nhau).

Ví dụ: Có tài liệu về số lao động của doanh nghiệp X vào các ngày trong quý I của năm nghiên cứu:

Bảng 6.3: Dãy số thời gian về số lao động

Ngày	1/1	15/1	20/2	13/3
Số lao động (người)	102	104	100	110

Các khoảng cách thời gian trong dãy số trên là không bằng nhau.

Muốn xây dựng một dãy số thời gian khoa học và chính xác, để phản ánh một cách đúng đắn sự phát triển của hiện tượng theo thời gian cần chú ý đảm bảo tính chất có thể so sánh được giữa các mức độ trong dãy số. Cụ thể là:

- Nội dung, phương pháp tính chỉ tiêu trong dãy số theo thời gian phải thống nhất.
- Phạm vi của hiện tượng nghiên cứu trước và sau phải nhất trí.
- Các khoảng cách thời gian trong dãy số nên bằng nhau (nhất là đối với dãy số thời kỳ).

Trong thực tế, do những nguyên nhân khác nhau, các yêu cầu trên có thể bị vi phạm, khi đó cần phải có sự chỉnh lý thích hợp để đảm bảo khả năng có thể so sánh được giữa các mức độ trong dãy số.

6.2. Các chỉ tiêu phân tích dãy số thời gian

6.2.1. Mức độ trung bình theo thời gian

Chỉ tiêu mức độ trung bình theo thời gian là số trung bình của các mức độ trong dãy số, phản ánh mức độ đại diện điển hình của dãy số thời gian. Tùy thuộc vào từng loại dãy số thời gian mà có các công thức tính khác nhau:

Đối với dãy số thời kỳ: Mức độ trung bình theo thời gian là số trung bình cộng giản đơn của các mức độ trong dãy số.

Công thức:
$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n} = \frac{\sum y_i}{n} \tag{6.1}$$

Trong đó:

y_i ($i = 1, 2, \dots, n$): Là các mức độ trong dãy số.

n : Số các mức độ trong dãy số.

Từ ví dụ ở bảng 6.1, ta có:

$$\bar{y} = \frac{200 + 240 + 276 + 300 + 320}{5} = 267,2 \text{ tỷ đồng.}$$

Như vậy, doanh thu trung bình một năm của doanh nghiệp trong khoảng thời gian từ 2011 đến 2015 là 267,2 tỷ đồng.

Đối với dãy số thời điểm:

Trường hợp dãy số thời điểm có khoảng cách thời gian bằng nhau:

Từ ví dụ bảng 6.2, ta có thể tính giá trị hàng hóa tồn kho trung bình của từng tháng trong quý I, dựa trên cơ sở số liệu đầu tháng và cuối tháng (tức là đầu tháng sau đó), với giả thiết giá trị hàng hóa tồn kho các ngày trong tháng biến động tương đối đều đặn.

Giá trị hàng hóa tồn kho bình quân của từng tháng như sau:

Tháng 1 là:
$$\bar{y}_1 = \frac{20 + 22}{2} = 21 \text{ tỷ đồng.}$$

Tháng 2 là: $\bar{y}_2 = \frac{22+28}{2} = 25$ tỷ đồng.

Tháng 3 là: $\bar{y}_3 = \frac{28+36}{2} = 32$ tỷ đồng.

Giá trị hàng hóa tồn kho trung bình trong quý I được tính bằng số trung bình cộng của giá trị hàng hóa tồn kho của các tháng 1, 2, 3:

$$\bar{y} = \frac{\bar{y}_1 + \bar{y}_2 + \bar{y}_3}{3}$$

$$\bar{y} = \frac{21 + 25 + 32}{3}$$

$$\text{Hay: } \bar{y} = \frac{\frac{20}{2} + 22 + 28 + \frac{36}{2}}{4-1} = 26 \text{ tỷ đồng.}$$

Như vậy, giá trị hàng hóa tồn kho trung bình trong quý I của doanh nghiệp là 26 tỷ đồng.

Từ đó, ta có công thức tổng quát tính mức độ trung bình theo thời gian từ dãy số thời điểm có khoảng cách thời gian bằng nhau:

$$\bar{y} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1} \quad (6.2)$$

Trong đó:

y_i ($i= 1,2,\dots,n$): Là các mức độ của dãy số thời điểm.

Trường hợp dãy số thời điểm có khoảng cách thời gian không bằng nhau:

$$\bar{y} = \frac{y_1 t_1 + y_2 t_2 + \dots + y_n t_n}{t_1 + t_2 + \dots + t_n} = \frac{\sum y_i t_i}{\sum t_i} \quad (6.3)$$

Trong đó:

y_i ($i= 1,2,\dots,n$): Là các mức độ của dãy số.

t_i ($i= 1,2,\dots,n$): Khoảng thời gian tương ứng với mức độ y_{ii} .

Ví dụ: Có tài liệu về số lao động của một doanh nghiệp trong tháng 9 như sau:

Ngày 1/9 có: 200 người.

Ngày 10/9 tuyển thêm 5 người.

Ngày 16/9 thêm 1 người.

Ngày 21/9 chuyển đi 3 người và từ đó đến cuối tháng 9 số lao động không thay đổi.

Ta lập bảng tính toán sau:

Bảng 6.4

Thời gian	Số ngày (t_i)	Số lao động (y_i)	($y_i t_i$)
Từ 1/9 đến 9/9	9	200	1800
Từ 10/9 đến 15/9	6	205	2130
Từ 16/9 đến 20/9	5	206	1030
Từ 21/9 đến 30/9	10	203	2030
Tổng số	30		6090

Số lao động trung bình trong tháng 9 của doanh nghiệp trên được tính theo công thức (6.3):

$$\bar{y} = \frac{\sum y_i t_i}{\sum t_i} = \frac{6090}{30} = 203 \text{ người.}$$

6.2.2. Lượng tăng (giảm) tuyệt đối

Lượng tăng (giảm) tuyệt đối là chỉ tiêu phản ánh sự thay đổi về mức độ tuyệt đối của hiện tượng giữa hai thời gian nghiên cứu. Nếu mức độ của hiện tượng tăng lên thì trị số của chỉ tiêu mang dấu dương (+) và ngược lại mang dấu (-).

Tùy theo mục đích nghiên cứu và gốc so sánh khi tính toán, có lượng tăng (giảm) tuyệt đối sau:

+ *Lượng tăng (giảm) tuyệt đối liên hoàn (hay từng kỳ):*

Là trị số chênh lệch giữa hai mức độ liên kế nhau trong dãy số hay chênh lệch giữa mức độ kỳ nghiên cứu với mức độ kỳ ngay trước đó (kỳ gốc liên hoàn). Chỉ tiêu này phản ánh sự biến động về mức độ tuyệt đối của hiện tượng giữa hai thời gian liên nhau.

Công thức tính: $\delta_i = y_i - y_{i-1}$ ($i = 2, 3, \dots, n$). (6.4)

Trong đó: δ_i : Lượng tăng (giảm) tuyệt đối liên hoàn ở thời gian thứ i .

y_i : Mức độ kỳ nghiên cứu i .

y_{i-1} : Mức độ kỳ gốc liên hoàn ($i-1$).

Theo số liệu ở bảng 6.1 có các lượng tăng (giảm) tuyệt đối liên hoàn sau:

Doanh thu năm 2012 so với năm 2011:

$$\delta_2 = y_2 - y_1 = 240 - 200 = 40 \text{ tỷ đồng.}$$

Doanh thu năm 2013 so với năm 2012:

$$\delta_3 = y_3 - y_2 = 276 - 240 = 36 \text{ tỷ đồng.}$$

Doanh thu năm 2014 so với năm 2013:

$$\delta_4 = y_4 - y_3 = 300 - 276 = 24 \text{ tỷ đồng.}$$

Doanh thu năm 2015 so với năm 2014:

$$\delta_5 = y_5 - y_4 = 320 - 300 = 20 \text{ tỷ đồng.}$$

Như vậy, doanh thu của doanh nghiệp có xu hướng tăng qua các năm.

+ *Lượng tăng (giảm) tuyệt đối định gốc (hay tính dồn):*

Là trị số chênh lệch giữa mức độ kỳ nghiên cứu và mức độ của kỳ nào đó được chọn làm gốc cố định cho mọi lần so sánh (thường là mức độ đầu tiên trong dãy số). Chỉ tiêu này phản ánh sự biến động về mức độ tuyệt đối của hiện tượng trong khoảng thời gian dài.

Công thức tính: $\Delta_i = y_i - y_1$ ($i = 2, 3, \dots, n$) (6.5)

Trong đó: Δ_i : Lượng tăng giảm tuyệt đối định gốc ở thời gian i .

y_1 : Mức độ kỳ gốc cố định.

Theo số liệu ở bảng 6.1, ta có các lượng tăng giảm tuyệt đối định gốc sau:

Doanh thu năm 2012 so với 2011:

$$\Delta_2 = y_2 - y_1 = 240 - 200 = 40 \text{ tỷ đồng.}$$

Doanh thu năm 2013 so với 2011:

$$\Delta_3 = y_3 - y_1 = 276 - 200 = 76 \text{ tỷ đồng.}$$

Doanh thu năm 2014 so với 2011:

$$\Delta_4 = y_4 - y_1 = 300 - 200 = 100 \text{ tỷ đồng.}$$

Doanh thu năm 2015 so với 2011:

$$\Delta_5 = y_5 - y_1 = 320 - 200 = 120 \text{ tỷ đồng.}$$

Mối quan hệ giữa lượng tăng (giảm) tuyệt đối liên hoàn và lượng tăng (giảm) tuyệt đối định gốc:

$$\Delta_n = y_n - y_1 = (y_2 - y_1) + (y_3 - y_2) + \dots + (y_n - y_{n-1}) = \sum \delta_i \quad (6.6)$$

Lượng tăng (giảm) tuyệt đối định gốc bằng tổng đại số của các lượng tăng (giảm) tuyệt đối liên hoàn.

+ *Lượng tăng (giảm) tuyệt đối trung bình:*

Là số trung bình cộng của các lượng tăng giảm tuyệt đối liên hoàn. Chỉ tiêu này phản ánh trung bình mỗi khoảng thời gian hiện tượng tăng hoặc giảm với mức độ tuyệt đối là bao nhiêu.

Công thức tính:
$$\bar{\delta} = \frac{\sum \delta_i}{n-1} = \frac{\Delta_n}{n-1} = \frac{y_n - y_1}{n-1} \quad (6.7)$$

Theo ví dụ trên:
$$\bar{\delta} = \frac{320 - 200}{5 - 1} = 30 \text{ tỷ đồng.}$$

Như vậy, trung bình một năm (trong khoảng thời gian từ năm 2011 đến 2015) doanh thu của doanh nghiệp tăng 30 tỷ đồng.

Từ công thức tính lượng tăng giảm tuyệt đối trung bình cho thấy: Chỉ nên tính trong trường hợp dãy số thời gian biến động theo một xu hướng nhất định.

6.2.3. Tốc độ phát triển

Tốc độ phát triển là số tương đối động thái (biểu hiện bằng số lần hay %) phản ánh xu hướng và trình độ phát triển của hiện tượng theo thời gian.

Tùy theo mục đích nghiên cứu và gốc so sánh, có các tốc độ phát triển sau:

+ *Tốc độ phát triển liên hoàn (hay từng kỳ):*

Là tỷ lệ so sánh giữa hai mức độ liên kế nhau trong dãy số hay giữa mức độ kỳ nghiên cứu với mức độ kỳ gốc liên hoàn. Chỉ tiêu này phản ánh xu hướng biến động và trình độ phát triển của hiện tượng nghiên cứu giữa hai thời gian liên nhau.

$$\text{Công thức tính: } t_i = \frac{y_i}{y_{i-1}} \quad (i = 2, 3, \dots, n) \quad (6.8)$$

Trong đó: t_i : Là tốc độ phát triển liên hoàn (biểu hiện bằng số lần hoặc %).

Theo số liệu ở bảng 6.1, ta có các tốc độ phát triển liên hoàn sau:

Doanh thu năm 2012 so với năm 2011:

$$t_2 = \frac{y_2}{y_1} = \frac{240}{200} = 1,2 \text{ lần hay } 120\%.$$

Doanh thu năm 2013 so với năm 2012:

$$t_3 = \frac{y_3}{y_2} = \frac{276}{240} = 1,15 \text{ lần hay } 115\%.$$

Doanh thu năm 2014 so với năm 2013:

$$t_4 = \frac{y_4}{y_3} = \frac{300}{276} = 1,087 \text{ hay } 108,7\%.$$

Doanh thu năm 2015 so với năm 2014:

$$t_5 = \frac{y_5}{y_4} = \frac{320}{300} = 1,0667 \text{ lần hay } 106,67\%$$

+ *Tốc độ phát triển định gốc (tính dồn):*

Là tỷ lệ so sánh giữa mức độ kỳ nghiên cứu với mức độ kỳ gốc cố định. Chỉ tiêu này phản ánh xu hướng biến động và trình độ phát triển của hiện tượng giữa kỳ nghiên cứu với một kỳ cố định.

Công thức tính: $T_i = \frac{y_i}{y_1}$ ($i= 2, 3, \dots, n$) (6.9)

Trong đó: T_i : Là tốc độ phát triển định gốc (biểu hiện bằng số lần hoặc %).

Theo số liệu ở bảng 6.1, ta có các tốc độ phát triển định gốc sau:

Doanh thu năm 2012 so với năm 2011:

$$T_2 = \frac{y_2}{y_1} = \frac{240}{200} = 1,2 \text{ lần hay } 120\%.$$

Doanh thu năm 2013 so với năm 2011:

$$T_3 = \frac{y_3}{y_2} = \frac{276}{200} = 1,38 \text{ lần hay } 138\%.$$

Doanh thu năm 2014 so với năm 2011:

$$T_4 = \frac{y_4}{y_3} = \frac{300}{200} = 1,5 \text{ lần hay } 150\%.$$

Doanh thu năm 2015 so với năm 2011:

$$T_5 = \frac{y_5}{y_4} = \frac{320}{200} = 1,6 \text{ lần hay } 160\%.$$

Mối quan hệ giữa tốc độ phát triển liên hoàn và tốc độ phát triển định gốc:

- Tốc độ phát triển định gốc của một thời kỳ bằng tích số của các tốc độ phát triển liên hoàn của thời kỳ đó:

$$T_n = t_2 t_3 \dots t_n = \prod t_i \quad (6.10)$$

- Thương của 2 tốc độ phát triển định gốc liên nhau bằng tốc độ phát triển liên hoàn giữa hai thời gian đó:

$$\frac{T_i}{T_{i-1}} = t_i \quad (i=2, 3, \dots, n) \quad (6.11)$$

+ *Tốc độ phát triển trung bình:*

Là số trung bình nhân của các tốc độ phát triển liên hoàn trong thời kỳ nghiên cứu. Chỉ tiêu này phản ánh trung bình giữa hai thời gian hiện tượng phát triển với tốc độ bao nhiêu lần hay bao nhiêu %.

$$\text{Công thức tính: } \bar{t} = \sqrt[n]{t_2 t_3 \dots t_n} = \sqrt[n]{T_n} = \sqrt[n]{\frac{y_n}{y_1}} \quad (6.12)$$

$$\text{Theo ví dụ trên ta có: } \bar{t} = \sqrt[5]{\frac{y_n}{y_1}} = \sqrt[5]{\frac{320}{200}} = 1,1247 \text{ lần hay } 112,47\%.$$

Tức là tốc độ phát triển trung bình hàng năm về doanh thu của doanh nghiệp là 112,47%.

Từ công thức tính tốc độ phát triển trung bình cho thấy: Chỉ nên tính trong trường hợp dãy số thời gian biến động theo một xu hướng nhất định.

6.2.4. Tốc độ tăng (giảm)

Chỉ tiêu này phản ánh mức độ của hiện tượng giữa hai thời gian đã tăng (giảm) bao nhiêu lần hoặc bao nhiêu %. Tương ứng với các tốc độ phát triển có các tốc độ tăng (giảm) sau:

+ *Tốc độ tăng (giảm) liên hoàn:*

Là tỷ lệ so sánh giữa lượng tăng giảm tuyệt đối liên hoàn với mức độ kỳ gốc liên hoàn. Chỉ tiêu này phản ánh giữa hai thời gian liên nhau hiện tượng đã tăng (giảm) bao nhiêu lần hoặc bao nhiêu %.

$$\text{Công thức tính: } a_i = \frac{\delta_i}{y_{i-1}} \text{ hoặc } a_i = t_i - 1 \text{ (hay } 100\%) \quad (6.13)$$

Trong đó: a_i là các tốc độ tăng giảm liên hoàn (biểu hiện bằng số lần hoặc %).

Từ các kết quả đã tính toán ở trên, ta có các tốc độ tăng qua các năm:

$$a_2 = t_2 - 1 = 1,2 - 1 = 0,2 \text{ lần hay } 20\%.$$

$$a_3 = t_3 - 1 = 1,15 - 1 = 0,15 \text{ lần hay } 15\%.$$

$$a_4 = t_4 - 1 = 1,087 - 1 = 0,087 \text{ lần hay } 8,7\%.$$

$$a_5 = t_5 - 1 = 1,0667 - 1 = 0,0667 \text{ lần hay } 6,67\%.$$

+ *Tốc độ tăng (giảm) định gốc:*

Là tỷ lệ so sánh giữa lượng tăng (giảm) tuyệt đối định gốc với mức độ kỳ gốc cố định. Chỉ tiêu này phản ánh hiện tượng ở kỳ nghiên cứu đã tăng (giảm) bao nhiêu lần hoặc bao nhiêu % so với kỳ gốc cố định.

Công thức tính: $A_i = \frac{\Delta_i}{y_1} = T_i - 1 \text{ (hay } 100\%)$ (6.14)

A_i : Tốc độ tăng (giảm) định gốc (biểu hiện bằng số lần hoặc %).

Tính được các tốc độ tăng (giảm) định gốc về doanh thu ở ví dụ trên:

$$A_2 = T_2 - 1 = 1,2 - 1 = 0,2 \text{ lần hay } 20\%.$$

$$A_3 = T_3 - 1 = 1,38 - 1 = 0,38 \text{ lần hay } 38\%.$$

$$A_4 = T_4 - 1 = 1,5 - 1 = 0,5 \text{ lần hay } 50\%.$$

$$A_5 = T_5 - 1 = 1,6 - 1 = 0,6 \text{ lần hay } 60\%.$$

+ *Tốc độ tăng (giảm) trung bình :*

Là chỉ tiêu phản ánh tốc độ tăng (giảm) đại diện của hiện tượng trong suốt thời gian nghiên cứu.

Công thức tính: $\bar{a} = \bar{t} - 1$ hoặc $\bar{a} = \bar{t}(\%) - 100$ (6.15)

Trong đó $\bar{a}$: Là tốc độ tăng (giảm) trung bình.

Theo kết quả tính toán trên, ta có: $\bar{a} = 1,1247 - 1 = 0,1247$ lần hay 12,47%.

Nghĩa là tốc độ tăng trung bình hàng năm về doanh thu của doanh nghiệp là 12,47%.

6.2.5. Giá trị tuyệt đối 1% tăng (giảm)

Là lượng tăng (giảm) tuyệt đối tương ứng với 1% của tốc độ tăng hoặc giảm liên hoàn.

$$\text{Công thức tính: } g_i = \frac{\delta_i}{a_i(\%)} = \frac{\delta_i}{\frac{\delta_i}{y_{i-1}} \times 100} = \frac{y_{i-1}}{100} \quad (6.16)$$

Trong đó g_i : Là giá trị tuyệt đối của 1% tăng (giảm).

Theo số liệu ở bảng 6.1 ta có:

$$g_2 = \frac{200}{100} = 2 \text{ tỷ đồng}$$

Tức là năm 2012 so với năm 2011 cứ 1% tăng của doanh thu thì tương ứng với mức tăng tuyệt đối là 2 tỷ đồng.

Tương tự tính cho các năm tiếp theo:

$$g_3 = \frac{240}{100} = 2,4 \text{ tỷ đồng}$$

$$g_4 = \frac{276}{100} = 2,76 \text{ tỷ đồng.}$$

$$g_5 = \frac{300}{100} = 3 \text{ tỷ đồng.}$$

Chỉ tiêu giá trị tuyệt đối của 1% tăng (giảm) không tính đối với tốc độ tăng (giảm) định gốc vì nó là số không đổi và bằng $\frac{y_1}{100}$.

Có thể trình bày kết quả tính toán các chỉ tiêu phân tích dãy số thời gian về doanh thu của doanh nghiệp trên bằng bảng thống kê sau:

Bảng 6.5: Kết quả tính toán các chỉ tiêu phân tích dãy số thời gian

Năm	Doanh thu (tỷ.đ)	Lượng tăng (giảm) tuyệt đối (tỷ.đ)		Tốc độ phát triển (%)		Tốc độ tăng (giảm) (%)		Giá trị tuyệt đối (1%) tăng (giảm)
	(y _i)	L.H(Δ _i)	Đ.G(Δ _i)	L.H(t _i)	Đ.G(T _i)	L.H(a _i)	Đ.G(A _i)	(g _i)
2011	200	-	-	-	-	-	-	-
2012	240	40	40	120	120	20	20	2
2013	276	36	76	115	138	15	38	2,4
2014	300	24	100	108,7	150	8,7	50	2,76
2015	320	20	120	106,67	160	6,67	60	3,0
TB	267,2	30		112,47		12,47		

Trên đây là 5 chỉ tiêu sử dụng để phân tích đặc điểm biến động của hiện tượng theo thời gian. Mỗi chỉ tiêu có ý nghĩa riêng, song giữa chúng có mối liên hệ với nhau nhằm giúp cho việc phân tích sâu sắc hơn.

6.3. Các phương pháp biểu hiện xu hướng phát triển cơ bản của hiện tượng

Sự biến động của hiện tượng qua thời gian thường chịu tác động của nhiều nhân tố. Có các nhân tố chủ yếu sẽ quyết định xu hướng phát triển cơ bản của hiện tượng. Bên cạnh đó, còn có những nhân tố ngẫu nhiên làm cho hiện tượng biến động sai lệch khỏi xu hướng đó. Việc xác định xu hướng biến động của hiện tượng có ý nghĩa quan trọng trong nghiên cứu thống kê. Vì vậy, cần thiết sử dụng các phương pháp thích hợp nhằm loại bỏ tác động ảnh hưởng của những nhân tố ngẫu nhiên, qua đó nêu lên xu hướng và tính quy luật phát triển của hiện tượng. Sau đây trình bày một số phương pháp thường sử dụng để biểu hiện xu hướng biến động cơ bản của hiện tượng theo thời gian.

6.3.1. Phương pháp mở rộng khoảng cách thời gian

Phương pháp này được áp dụng khi dãy số thời kỳ có nhiều mức độ và các khoảng cách thời gian giữa các mức độ ngắn, đồng thời chưa thấy rõ xu hướng phát triển cơ bản của hiện tượng. Trên cơ sở dãy số đó, có thể xây dựng một dãy số mới với các khoảng cách thời gian dài hơn, khi đó sẽ rút bớt được số lượng các mức độ trong dãy số. Bằng phương pháp mở rộng khoảng cách thời gian như biến đổi mức độ hàng ngày thành mức độ hàng tháng, từ tháng thành quý, từ quý thành năm...

Ví dụ: Có tài liệu về giá trị sản xuất hàng tháng trong năm của một doanh nghiệp như sau:

Bảng 6.6: Dãy số thời gian về giá trị sản xuất

Tháng	Giá trị sản xuất GO (Tỷ.đ)	Tháng	Giá trị sản xuất GO (Tỷ.đ)
1	50	7	53
2	52	8	55
3	48	9	52
4	51	10	58
5	49	11	54
6	56	12	60

Dãy số trên cho thấy: giá trị sản xuất của doanh nghiệp (GO) qua các tháng khi tăng khi giảm, không phản ánh rõ xu hướng biến động. Ta có thể mở rộng khoảng cách thời gian từ tháng sang quý. Kết quả sẽ được dãy số thời gian mới với khoảng cách thời gian theo quý như sau:

Bảng 6.7:

Quý	Giá trị sản xuất GO (tỷ.đ)
I	150
II	156
III	160
IV	172

Do khoảng cách thời gian được mở rộng, nên trong mỗi mức độ của dãy số mới thì sự tác động của các nhân tố ngẫu nhiên phần nào được bù trừ, triệt tiêu và qua đó thấy rõ hơn xu hướng biến động của hiện tượng. Theo kết quả bảng (6.7) thì giá trị sản xuất của doanh nghiệp tăng dần qua các quý ở trong năm nghiên cứu.

6.3.2. Phương pháp số trung bình di động (số trung bình trượt)

Phương pháp này được vận dụng trong trường hợp các mức độ trong dãy số có biến động ngẫu nhiên nhưng mức biến động không lớn.

Số trung bình di động là số trung bình cộng của một nhóm nhất định các mức độ trong dãy số, được tính bằng cách loại trừ dần mức độ đầu, đồng thời thêm vào đó mức độ tiếp theo sao cho số lượng các mức độ tham gia tính số trung bình di động không thay đổi.

Giả sử có dãy số thời gian: $y_1, y_2, \dots, y_{n-1}, y_n$. Nếu tính số trung bình di động của một nhóm 3 mức độ, ta sẽ có các số trung bình di động sau:

$$\overline{y_1} = \frac{y_1 + y_2 + y_3}{3}$$

$$\overline{y_2} = \frac{y_2 + y_3 + y_4}{3}$$

.....

$$\overline{y_{n-2}} = \frac{y_{n-2} + y_{n-1} + y_n}{3}$$

Dãy số mới là dãy số mà các mức độ sẽ là các số trung bình di động. Số trung bình di động có tác dụng san bằng ảnh hưởng của những nhân tố ngẫu nhiên đồng thời làm giảm các mức độ trong dãy số mới.

Ví dụ: Có tài liệu về kết quả sản xuất của một công ty qua 8 năm.

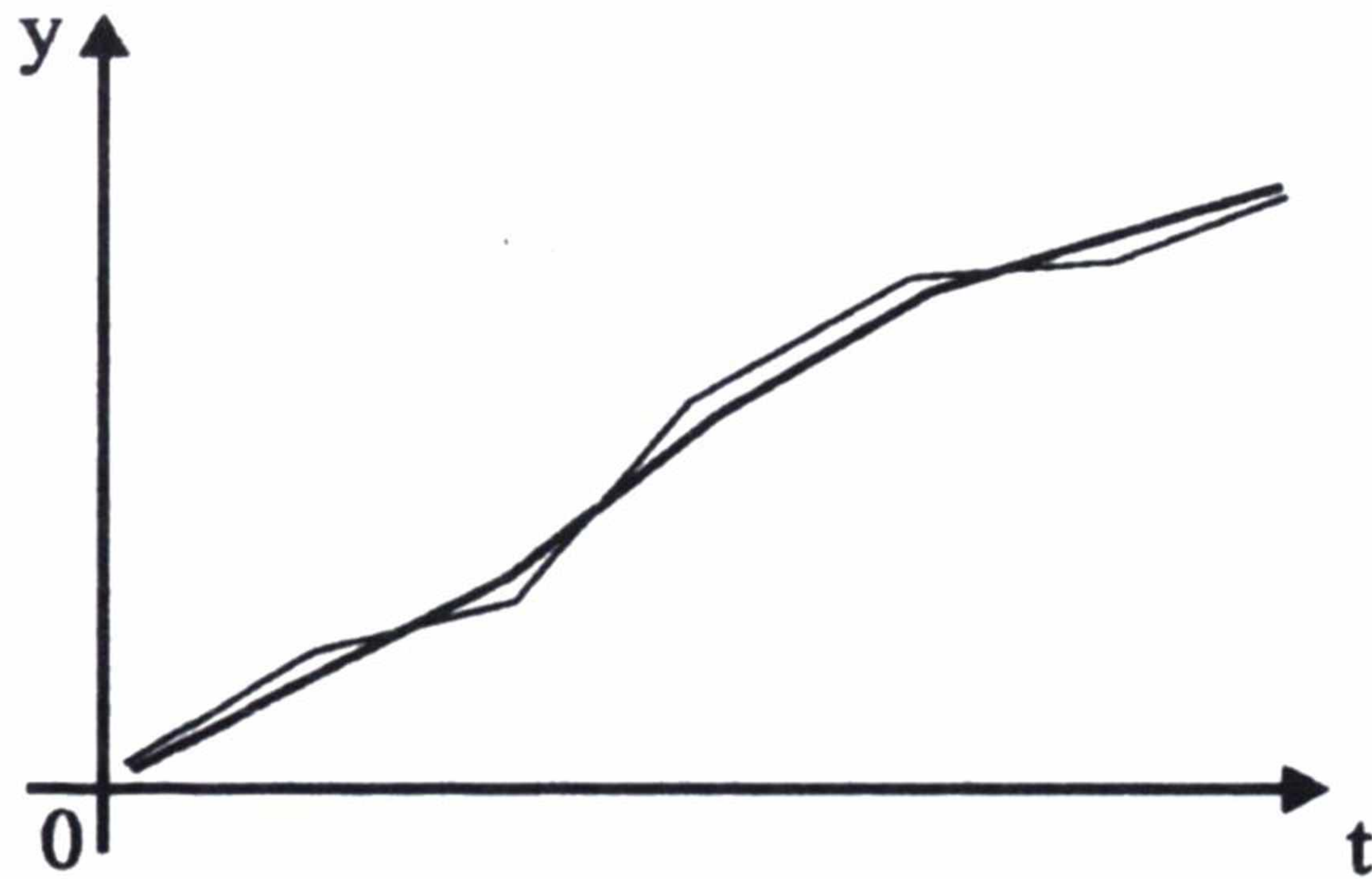
Nếu tính số trung bình di động cho nhóm 3 mức độ, ta sẽ có kết quả sau:

Bảng 6.8: Dãy số thời gian và các số trung bình di động

Năm	Sản lượng (tấn) (y_i)	Số trung bình di động ($\bar{y}_i$)
2008	200	-
2009	212	207
2010	209	219
2011	236	221
2012	218	233
2013	245	234
2014	239	238
2015	230	-

Có thể biểu hiện sự biến động của hai dãy số này trên đồ thị và so sánh, sẽ cho thấy đồ thị của dãy số trung bình di động ($\bar{y}_i$) phẳng hơn đồ thị của dãy số thực tế (y_i) vì số trung bình di động đã san bằng phần nào ảnh hưởng của các nhân tố ngẫu nhiên.

Đồ thị 6.1



Trong đó: — Đồ thị dãy số thực tế
 — Đồ thị dãy số trung bình di động

7

Vấn đề là việc chọn nhóm bao nhiêu mức độ để tính số trung bình di động cho hợp lý. Trước hết, phải dựa vào đặc điểm biến động của hiện tượng và số lượng các mức độ của dãy số thời gian. Nếu sự biến động của hiện tượng theo chu kỳ, thì chọn số mức độ theo đúng chu kỳ đó; hiện tượng biến động tương đối đều đặn và số lượng các mức độ của dãy số không nhiều thì tính số trung bình di động từ ba mức độ. Nếu hiện tượng biến động lớn, dãy số có nhiều mức độ thì có thể tính số trung bình di động từ bốn, năm, ... mức độ. Số trung bình di động càng tính từ nhóm nhiều mức độ càng có tác dụng san bằng ảnh hưởng của các nhân tố ngẫu nhiên, nhưng khi đó số lượng các mức độ của dãy số trung bình di động lại càng giảm, điều đó ảnh hưởng tới việc biểu hiện xu hướng phát triển cơ bản của hiện tượng.

6.3.3. Phương pháp hồi quy

Phương pháp này được áp dụng trong trường hợp dãy số thời gian có nhiều biến động lớn, khi tăng khi giảm thất thường.

Trên cơ sở dãy số thời gian, lựa chọn một dạng phương trình thích hợp để biểu hiện xu hướng biến động của hiện tượng theo thời gian (gọi là phương trình hồi quy hay hàm xu thế). Dạng tổng quát của phương trình hồi quy là:

$$\hat{y}_t = f(t, a_0, a_1, \dots, a_n) \quad (6.17)$$

Trong đó: $\hat{y}_t$: Là mức độ lý thuyết hay trị số điều chỉnh của dãy số thời gian.

t : Thứ tự thời gian.

$a_0, a_1, \dots, a_n$: Các tham số của phương trình.

Vấn đề có tính chất quyết định là việc lựa chọn đúng đắn dạng của phương trình hồi quy phản ánh sự biến động của hiện tượng. Để giải quyết vấn đề này đòi hỏi phải dựa vào việc phân tích đặc điểm biến động của hiện tượng qua thời gian, đồng thời kết hợp với phương pháp đồ thị và một số tiêu chuẩn khác như sai số chuẩn của mô hình.

Sau đây là một số dạng phương trình hồi quy đơn giản thường được sử dụng:

+ *Phương trình hồi quy tuyến tính*:

Trong trường hợp hiện tượng phát triển tăng (giảm) tương đối đều đặn theo một chiều hướng nhất định, tức là khi dãy số có các lượng tăng (giảm) tuyệt đối liên hoàn xấp xỉ bằng nhau, có thể chọn phương trình đường thẳng có dạng:

$$\hat{y}_t = a_0 + a_1 t \quad (6.18)$$

Hệ phương trình để xác định các hệ số a_0, a_1 :

$$\begin{cases} na_0 + a_1 \cdot \sum t = \sum y \\ a_0 \cdot \sum t + a_1 \cdot \sum t^2 = \sum t \cdot y \end{cases} \quad (6.19)$$

+ *Phương trình pa-ra-bôn*:

Trường hợp các mức độ của dãy số thời gian tăng dần theo thời gian, đạt giá trị cực đại sau đó lại giảm dần; hoặc ngược lại giảm dần theo thời gian, đạt giá trị cực tiểu rồi lại tăng dần.

Phương trình hồi quy:

$$\hat{y}_t = a_0 + a_1 t + a_2 t^2 \quad (6.20)$$

Hệ phương trình xác định các hệ số a_0, a_1, a_2 :

$$\begin{cases} na_0 + a_1 \cdot \sum t + a_2 \sum t^2 = \sum y \\ a_0 \cdot \sum t + a_1 \sum t^2 + a_2 \sum t^3 = \sum ty \\ a_0 \cdot \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 = \sum t^2 y \end{cases} \quad (6.21)$$

+ *Phương trình hy-pe-bôn*:

Trường hợp các mức độ của dãy số thời gian có xu hướng giảm dần theo thời gian, nhưng không đều nhau. Dạng phương trình hy-pe-bôn như sau:

$$\hat{y}_x = a_0 + \frac{a_1}{t} \quad (6.22)$$

Hệ phương trình để xác định các hệ số a_0, a_1 :

$$\begin{cases} na_0 + a_1 \sum \frac{1}{t} = \sum y \\ a_0 \sum \frac{1}{t} + a_1 \sum \frac{1}{t^2} = \sum \frac{y}{t} \end{cases} \quad (6.23)$$

+ *Phương trình hàm số mũ*:

Trường hợp dãy số thời gian có các tốc độ phát triển liên hoàn xấp xỉ nhau, phương trình hồi quy có dạng:

$$\hat{y}_t = a_0 \cdot a_1^t \quad (6.24)$$

Hệ phương trình để xác định các hệ số a_0, a_1 :

$$\begin{cases} n \ln a_0 + \ln a_1 \sum t = \sum \ln y \\ \ln a_0 \sum t + \ln a_1 \sum t^2 = \sum \ln yt \end{cases} \quad (6.25)$$

Lưu ý: Các tham số a_i ($i=0,1,2, \dots, n$) của các phương trình hồi quy trên thường được xác định bằng phương pháp bình phương nhỏ nhất: $\sum (y_t - \bar{y}_t)^2 = \min$ (đã được trình bày ở chương 5).

Ta xét ví dụ: Trở lại số liệu ở bảng (6.1), ta thấy doanh thu của doanh nghiệp qua các năm tăng lên, có thể sử dụng phương trình đường thẳng để biểu hiện xu hướng biến động của doanh thu theo thời gian.

Phương trình có dạng:

$$\bar{y}_t = a_0 + a_1 t$$

Xác định giá trị các hệ số a_0, a_1 từ hệ phương trình sau:

$$\begin{cases} na_0 + a_1 \cdot \sum t = \sum y \\ a_0 \cdot \sum t + a_1 \cdot \sum t^2 = \sum t \cdot y \end{cases}$$

Ta lập bảng tính toán sau:

Bảng 6.9: Các đại lượng để tính các tham số

Năm	Doanh thu (y)	Phần tính toán			
		t	t.y	t ²	$\hat{y}$
2011	200	1	200	1	207,2
2012	240	2	480	4	237,2
2013	276	3	828	9	267,2
2014	300	4	1200	16	297,2
2015	320	5	1600	25	327,2
Tổng	1336	15	4308	55	

Thay số liệu vào hệ phương trình, ta có:

$$\begin{cases} 5a_0 + 15a_1 = 1336 \\ 15a_0 + 55a_1 = 4308 \end{cases}$$

Giải hệ phương trình được: $a_0 = 177,2; a_1 = 30$.

Phương trình hồi quy: $\hat{y}_t = 177,2 + 30 t$

Như vậy, hệ số hồi quy: $a_1 = 30$ phản ánh mức tăng trung bình một năm của doanh thu là 30 tỷ đồng.

Từ phương trình hồi quy đã xác định, có thể tính các giá trị lý thuyết ($\hat{y}$) của dãy số thời gian, bằng cách lần lượt thay các giá trị $t = 1, 2, 3, \dots$ vào phương trình (kết quả tính toán trên cột của bảng 6.9).

6.3.4. Phương pháp biểu hiện biến động thời vụ

Trong thực tế, sự biến động của một số hiện tượng kinh tế - xã hội thường có tính chất thời vụ. Nghĩa là: Hàng năm, trong từng thời gian nhất định, sự biến động của hiện tượng được lặp đi lặp lại. Nguyên nhân gây ra biến động thời vụ là do ảnh hưởng của điều kiện tự nhiên (thời tiết, khí hậu) và phong tục tập quán sinh hoạt của dân cư. Biến động thời vụ thường biểu hiện rõ nhất trong sản xuất nông nghiệp. Trong các ngành khác như công nghiệp, xây dựng, thương mại, dịch vụ... cũng đều có những biến động thời vụ.

Biến động thời vụ của hiện tượng có ảnh hưởng lớn tới sản xuất kinh doanh, tới đời sống sinh hoạt, làm một số ngành khi thì nhộn nhịp, khẩn trương; khi thì nhàn rỗi và bị thu hẹp lại.

Nghiên cứu biến động thời vụ giúp các nhà quản lý đề ra những chủ trương biện pháp phù hợp, kịp thời nhằm hạn chế những ảnh hưởng của biến động thời vụ đối với sản xuất kinh doanh và sinh hoạt của xã hội.

Nghiên cứu thống kê thường dựa vào nguồn số liệu về hiện tượng qua nhiều năm, cụ thể qua các tháng hoặc các quý (ít nhất là trong 3 năm). Sử dụng phương pháp chỉ số thời vụ để xác định tính chất và mức độ của biến động thời vụ.

Trường hợp biến động thời vụ qua những thời gian nhất định của các năm tương đối ổn định, không có hiện tượng tăng hoặc giảm rõ rệt thì chỉ số thời vụ có thể tính theo công thức sau:

$$I_i = \frac{\overline{y_i}}{\overline{y_0}} \quad (6.26)$$

Trong đó:

I_i : Chỉ số thời vụ của tháng hoặc quý thứ i .

$\overline{y_i}$: Mức độ trung bình của từng tháng hoặc quý trong các năm nghiên cứu.

$\overline{y_0}$: Mức độ trung bình chung của một tháng hoặc quý trong các năm nghiên cứu.

Chỉ số thời vụ được tính bằng số lần hoặc %. Nếu $I_i > 1$ (100%), thì hiện tượng biến động tăng vào thời gian i ; ngược lại nếu $I_i < 1$ (100%), thì hiện tượng biến động giảm.

Ví dụ: Có tài liệu về doanh thu tiêu thụ một loại hàng hóa qua các quý của 3 năm tại doanh nghiệp A như sau:

Bảng 6.10: Doanh thu tiêu thụ hàng hóa trong 3 năm

Năm Quý	Doanh thu (Tr.đ)			$\bar{y}_i$	$I_i = \frac{\bar{y}_i}{y_0} 100\%$
	2013	2014	2015		
I	520	532	538	530	106
II	450	480	465	465	93
III	300	315	360	325	65
IV	650	703	687	680	136
Cả năm	1920	2030	2050	500	

Từ tài liệu về doanh thu của các quý qua 3 năm, muốn tính chỉ số thời vụ, trước hết tính doanh thu trung bình của từng quý trong cả 3 năm ($\bar{y}_i$):

Quý I: $\bar{y}_1 = \frac{520 + 532 + 538}{3} = 530$ triệu đồng

Quý II: $\bar{y}_2 = \frac{450 + 480 + 465}{3} = 465$ triệu đồng

Quý III: $\bar{y}_3 = \frac{300 + 315 + 360}{3} = 325$ triệu đồng

Quý IV: $\bar{y}_4 = \frac{650 + 703 + 687}{3} = 680$ triệu đồng

Doanh thu trung bình chung một quý trong cả 3 năm là:

$\bar{y}_0 = \frac{1920 + 2030 + 2050}{12} = 500$ triệu đồng

Chỉ số thời vụ của từng quý:

$$\text{Quý I: } I_1 = \frac{530}{500} = 1,06 \text{ lần hay } 106\%$$

$$\text{Quý II: } I_2 = \frac{465}{500} = 0,93 \text{ lần hay } 93\%$$

$$\text{Quý III: } I_3 = \frac{325}{500} = 0,65 \text{ lần hay } 65\%$$

$$\text{Quý IV: } I_4 = \frac{680}{500} = 1,36 \text{ lần hay } 136\%$$

Như vậy, doanh thu tiêu thụ mặt hàng tăng mạnh vào quý IV, rồi đến quý I; giảm mạnh vào quý III, rồi đến quý II.

Trường hợp biến động thời vụ qua những thời gian nhất định của các năm có sự tăng (giảm) rõ rệt thì chỉ số thời vụ có thể được tính theo công thức sau:

$$I_i = \frac{\sum \frac{\bar{y}_i}{y} \times 100}{n} \quad (6.27)$$

Trong đó: y_i : Mức độ thực tế.

$\bar{y}_i$: Mức độ lý thuyết tính được bằng phương trình hồi quy.

n : Số năm nghiên cứu.

6.4. Dự báo thống kê

6.4.1. Khái niệm và phân loại dự báo

Dự báo hiểu theo nghĩa chung nhất là xác định mức độ, trạng thái của hiện tượng nghiên cứu trong tương lai. Ngày nay dự báo là nhu cầu không thể thiếu được trong mọi hoạt động kinh tế - xã hội, khoa học - kỹ thuật, được sử dụng rộng rãi trong các lĩnh vực với nhiều loại và phương pháp dự báo khác nhau.

- Căn cứ theo thời gian dự báo có thể phân loại dự báo thành ba loại:

Dự báo ngắn hạn: Là những dự báo có thời gian dưới ba năm, thường phục vụ cho lập kế hoạch ngắn hạn, là công cụ quan trọng để tổ chức quản lý thường xuyên các hoạt động của các đơn vị, các cấp, các ngành.

Dự báo trung hạn: Là những dự báo có thời hạn từ ba đến năm năm, thường phục vụ cho việc lập kế hoạch trung hạn.

Dự báo dài hạn: Là những dự báo có thời hạn từ 5 năm trở lên, thường dự báo các mục tiêu chiến lược về kinh tế - xã hội phục vụ cho xây dựng kế hoạch định hướng phát triển dài hạn.

Tuy nhiên, cách phân loại này chỉ mang tính tương đối, tùy thuộc vào từng loại hiện tượng và chỉ tiêu dự báo mà quy định thời gian dự báo cho phù hợp.

- Căn cứ vào phương pháp dự báo có thể chia thành ba nhóm:

Dự báo bằng phương pháp chuyên gia: Là dự báo tiến hành trên cơ sở tổng hợp, xử lý, phân tích những ý kiến của các chuyên gia. Phương pháp này có ưu thế trong trường hợp dự báo những hiện tượng có tầm bao quát rộng, phức tạp, chịu sự chi phối của nhiều yếu tố. Ví dụ như dự báo về khoa học - kỹ thuật, sự thay đổi về môi trường trong thời gian dài.

Dự báo theo mô hình hồi quy: Mức độ cần dự báo phải được xây dựng trên cơ sở mô hình hồi quy phù hợp với đặc điểm và xu thế phát triển của hiện tượng. Để xây dựng mô hình hồi quy, đòi hỏi phải có tài liệu về hiện tượng cần dự báo và các hiện tượng khác có liên quan. Phương pháp dự báo này thường dùng để dự báo trung hạn và dài hạn ở tầm vĩ mô.

Dự báo dựa vào dãy số thời gian: Trên cơ sở dãy số thời gian phản ánh xu hướng biến động của hiện tượng trong thời gian đã qua để xác định mức độ của hiện tượng trong tương lai.

Ngoài ra, còn có thể phân loại dự báo theo đối tượng như dự báo khoa học, dự báo kinh tế, dự báo xã hội, dự báo tự nhiên...

6.4.2. Dự báo thống kê

Dự báo thống kê là trên cơ sở tài liệu thống kê về các hiện tượng nghiên cứu trong thời gian đã qua, sử dụng các phương pháp thích hợp để tính toán mức độ tương lai của hiện tượng.

Tài liệu sử dụng trong dự báo thống kê thường là dãy số thời gian về hiện tượng. Việc sử dụng dãy số thời gian trong dự báo có những ưu điểm cơ bản sau đây:

- + Dãy số thời gian sử dụng cho dự báo không đòi hỏi có quá nhiều mức độ như dự báo dựa vào mô hình hồi quy.

- + Việc xây dựng mô hình dự báo được tiến hành tương đối đơn giản, ít bị ràng buộc bởi các giả thiết như trong xây dựng mô hình hồi quy.

- + Dự báo dựa vào dãy số thời gian sẽ rất thuận lợi cho việc ứng dụng tin học. Nhờ đó việc tính toán trở nên thuận tiện, đồng thời cho phép lựa chọn mô hình dự báo phù hợp.

Tuy nhiên để kết quả dự báo có độ tin cậy cao, đòi hỏi dãy số thời gian sử dụng cho dự báo phải đảm bảo khoa học và chính xác, các mức độ của dãy số phải so sánh được với nhau. Mặt khác số lượng các mức độ của dãy số phải đủ lớn, không quá nhiều cũng không quá ít.

Việc xác định số lượng các mức độ của dãy số thời gian sử dụng cho dự báo phải dựa trên cơ sở phân tích đặc điểm biến động của hiện tượng. Nếu biến động của hiện tượng qua thời gian có tính chất tương đối ổn định, thường sử dụng dãy số từ 7 đến 10 mức độ. Trường hợp hiện tượng có biến động đột xuất ở thời gian cuối do các nhân tố mới xuất hiện, làm thay đổi xu hướng phát triển của hiện tượng thì có thể sử dụng một số mức độ cuối của dãy số thời gian để xây dựng mô hình dự báo.

Có thể sử dụng nhiều phương pháp trong dự báo thống kê. Trong đó một số phương pháp dựa trên các công thức tính toán thống kê như dự báo dựa vào lượng tăng (giảm) trung bình, tốc độ phát triển trung bình, dựa vào hàm xu thế,... Một số phương pháp dự báo được vận dụng dựa vào các kết quả của toán học như dự báo bằng phương pháp san bằng mũ, bằng mô hình tuyến tính ngẫu nhiên,...

Sau đây là trình bày một số phương pháp dự báo thống kê ngắn hạn dựa trên cơ sở dãy số thời gian thường được sử dụng.

6.4.3. Một số phương pháp dự báo thống kê thông dụng

- *Dự báo dựa vào lượng tăng (giảm) tuyệt đối trung bình:*

Mô hình dự báo:

$$\hat{Y}_{n+L} = y_n + \bar{\delta} L \tag{6.28}$$

Trong đó:

- $\hat{y}_{n+L}$: Mức độ dự báo của thời gian $n+L$.
- y_n : Mức độ cuối cùng của dãy số thời gian.
- $\bar{\delta}$: Lượng tăng (giảm) tuyệt đối trung bình.
- L : Tầm xa của dự báo (1, 2, 3,...).

Phương pháp này được áp dụng trong trường hợp dự báo đối với hiện tượng nghiên cứu phát triển tương đối ổn định, dãy số thời gian có các lượng tăng (giảm) tuyệt đối liên hoàn xấp xỉ nhau.

Ví dụ: Trở lại ví dụ ở bảng (6.1), ta tính được lượng tăng tuyệt đối trung bình một năm về doanh thu của doanh nghiệp giai đoạn 2011 đến 2015 là:

$$\bar{\delta} = \frac{320 - 200}{5 - 1} = 30 \text{ tỷ đồng.}$$

Dự báo doanh thu của doanh nghiệp cho năm 2016 với $L = 1$:

$$\hat{Y}_{2013} = 320 + 30.1 = 360 \text{ tỷ đồng.}$$

Dự báo doanh thu của doanh nghiệp cho năm 2017 với $L = 2$:

$$\hat{Y}_{2014} = 320 + 30.2 = 390 \text{ tỷ đồng.}$$

- *Dự báo dựa vào tốc độ phát triển trung bình:*

Mô hình dự báo:

$$\hat{y}_{n+L} = y_n (\bar{t})^L \tag{6.29}$$

Trong đó: $\bar{t}$ là tốc độ phát triển trung bình.

Phương pháp này được áp dụng trong trường hợp dự báo đối với hiện tượng nghiên cứu có nhịp độ phát triển đều đặn, dãy số thời gian có các tốc độ phát triển liên hoàn xấp xỉ nhau.

Ví dụ: Trở lại bảng (6.1), tính được tốc độ phát triển trung bình một năm về doanh thu của doanh nghiệp giai đoạn 2011 đến 2015 là:

$$\bar{t} = \sqrt[n-1]{\frac{y_n}{y_1}} = \sqrt[5-1]{\frac{320}{200}} = 1,1247 \text{ lần hay } 112,47\%.$$

Dự báo doanh thu của doanh nghiệp cho năm 2016 là:

$$\hat{Y}_{2016} = 320.(1,1247)^1 = 359,9 \text{ tỷ đồng}$$

Dự báo doanh thu của doanh nghiệp cho năm 2017 là:

$$\hat{Y}_{2017} = 320.(1,1247)^2 = 404,78 \text{ tỷ đồng}$$

• *Dự báo dựa vào hàm xu thế:*

Sau khi xác định được hàm xu thế biểu hiện xu hướng biến động của hiện tượng theo thời gian, có thể dựa vào đó để dự báo theo mô hình sau:

$$\hat{y}_{n+L} = f(t + L) \tag{6.30}$$

Ví dụ: Trở lại số liệu ở bảng (6.9), ta đã xác định được hàm xu thế (phương trình hồi quy) biểu hiện xu hướng phát triển doanh thu của doanh nghiệp là: $\hat{y}_t = 177,2 + 30 t$.

Dự báo doanh thu cho năm 2016 (với $t=6$) :

$$\hat{y}_t = 177,2 + 30.6 = 357,2 \text{ tỷ đồng}$$

Dự báo doanh thu cho năm 2017 (với $t=7$):

$$\hat{y}_t = 177,2 + 30.7 = 387,2 \text{ tỷ đồng}$$

Phần trên đã trình bày ba mô hình dự báo đều dựa trên cơ sở dãy số thời gian. Tuy nhiên, với mỗi mô hình lại cho các kết quả dự báo là khác nhau. Để lựa chọn mô hình dự báo nào cho kết quả dự báo tốt hơn tức là mức độ dự báo sát với mức độ thực tế hơn, có thể sử dụng một trong hai tiêu chuẩn sau:

- Tổng bình phương của sai số dự báo là nhỏ nhất:

$$SSE = \sum (y_t - \hat{y}_t)^2 \text{ min}$$

Trong đó: y_t : Mức độ thực tế ở thời gian t .

$\hat{y}_t$: Mức độ dự báo ở thời gian t .

- Sai số chuẩn của mô hình dự báo là nhỏ nhất:

$$SE = \sqrt{\frac{SSE}{n - p}} \text{ min}$$

Trong đó: p : Số lượng các tham số của mô hình dự báo.

n : Số lượng các mức độ của dãy số thời gian.

Ngoài ra, các mô hình dự báo thường được sử dụng trên đây có thể tiến hành dự báo thống kê bằng các phương pháp khác như: Dựa vào chỉ số thời vụ (nếu hiện tượng kinh tế - xã hội có biến động thời vụ), phương pháp san bằng mũ hay dự báo bằng một số mô hình tuyến tính ngẫu nhiên,...

Mức độ chính xác của kết quả dự báo phụ thuộc vào đặc điểm biến động của hiện tượng trong thời gian đã qua, độ dài thời gian của thời kỳ trước (số lượng các mức độ của dãy số thời gian) tầm xa dự báo và việc lựa chọn phương pháp dự báo.

TÓM TẮT CHƯƠNG 6

Hiện tượng kinh tế - xã hội không ngừng biến động theo thời gian. Thống kê sử dụng phương pháp dãy số thời gian để phân tích đặc điểm, xu hướng biến động của hiện tượng theo thời gian. Trên cơ sở đó, dự báo mức độ tương lai của hiện tượng. Căn cứ vào đặc điểm tồn tại của hiện tượng theo thời gian, có hai loại: Dãy số thời kỳ và dãy số thời điểm.

Để phân tích và đo lường sự biến động của hiện tượng qua thời gian, sử dụng các chỉ tiêu: Mức độ trung bình theo thời gian, lượng tăng (giảm) tuyệt đối, tốc độ phát triển, tốc độ tăng (giảm) và giá trị tuyệt đối của 1% tăng (giảm).

Hiện tượng biến động qua thời gian thường chịu ảnh hưởng của nhiều nhân tố. Để biểu hiện rõ quy luật phát triển của hiện tượng, có thể sử dụng các phương pháp: Mở rộng khoảng cách thời gian, số bình quân di động, phương pháp hồi quy, chỉ số thời vụ. Mỗi phương pháp được áp dụng trong các trường hợp khác nhau, tùy theo đặc điểm, tính chất của dãy số thời gian.

Dự báo (dự đoán) thống kê là xác định mức độ của hiện tượng trong tương lai trên cơ sở các tài liệu thống kê về hiện tượng. Tài liệu sử dụng trong dự báo thống kê thường là dãy số thời gian. Có thể sử dụng nhiều phương pháp trong dự báo thống kê tùy theo đặc điểm biến động của hiện tượng theo thời gian. Một số phương pháp dự báo thống kê đơn giản thường được sử dụng như: Dự báo dựa vào lượng tăng (giảm) tuyệt đối trung bình, dự báo dựa vào tốc độ phát triển trung bình, dự báo dựa vào hàm xu thế. Mức độ chính xác của kết quả dự báo phụ thuộc vào đặc điểm biến động của hiện tượng, số lượng các mức độ của dãy số thời gian và tầm xa dự báo.

CÂU HỎI ÔN TẬP CHƯƠNG 6

1. Trình bày khái niệm, ý nghĩa của dãy số thời gian. Các loại dãy số thời gian?
2. Phân tích các điều kiện xây dựng dãy số thời gian?
3. Ý nghĩa, phương pháp tính các chỉ tiêu phân tích dãy số thời gian?
4. Ý nghĩa nghiên cứu xu hướng phát triển cơ bản của hiện tượng theo thời gian? Trình bày các phương pháp nghiên cứu xu hướng phát triển cơ bản của hiện tượng?
5. Khái niệm, ý nghĩa và phân loại dự báo?
6. Trình bày các phương pháp dự báo thống kê thường dùng?

BÀI TẬP CHƯƠNG 6

Bài 1:

Giá trị hàng hoá tồn kho của Công ty X như sau:

Thời điểm	1/1	1/2	1/3	1/4
Chỉ tiêu				
Giá trị hàng hoá tồn kho (trđ)	4200	5100	5200	7100

Yêu cầu: Hãy tính giá trị hàng hoá tồn kho trung bình của công ty vào các thời gian:

- a) Từng tháng?
- b) Quý I (một tháng trong quý I)?

Bài 2:

Số lao động trong danh sách của công ty X như sau:

Ngày	Số người	Ngày	Số người
1/1	90	1/7	94
1/2	90	1/8	92
1/3	98	1/9	94
1/4	96	1/10	96
1/5	96	1/11	98
1/6	92	1/12	96

Biết số lao động trong danh sách của công ty ngày 31/12 là 94 người.

Hãy tính số lao động trong danh sách trung bình của công ty vào các thời gian:

- a) Mỗi tháng?
- b) Mỗi quý (một tháng trong quý)?
- c) 6 tháng đầu năm và 6 tháng cuối năm?
- d) Cả năm?

Bài 3:

Mức bán lẻ hàng hoá của một công ty như sau:

Chi tiêu	2011	2012	2013	2014	2015
Mức bán lẻ (tỷ đồng)	3000	3300	3600	4000	4800

Yêu cầu tính các chỉ tiêu:

- a) Mức độ trung bình theo thời gian?
- b) Lượng tăng tuyệt đối?
- c) Tốc độ phát triển?
- d) Tốc độ tăng?
- e) Giá trị tuyệt đối của 1% tăng?

Bài 4:

Có tài liệu về tình hình sản xuất của một doanh nghiệp như sau:

Chi tiêu	Tháng 1	Tháng 2	Tháng 3
- Giá trị sản xuất thực tế (triệu đồng)	31.620	33.600	33.800
- Tỷ lệ % hoàn thành kế hoạch giá trị sản xuất	102	105	104
- Số công nhân ngày đầu tháng (người)	300	304	304

Biết thêm số công nhân ngày 1/4 là 308 người.

Hãy tính:

- a) Giá trị sản xuất thực tế trung bình một tháng trong quý I?
- b) Số công nhân trung bình trong mỗi tháng và cả quý I?
- c) Năng suất lao động trung bình của công nhân mỗi tháng và 1 tháng trong quý I?
- d) Tỷ lệ % hoàn thành kế hoạch giá trị sản xuất trong quý I. Cho nhận xét về tình hình sản xuất của doanh nghiệp trên?

Bài 5a:

Kế hoạch 5 năm của một doanh nghiệp dự kiến tăng sản lượng 20%; kế hoạch này đã hoàn thành vượt mức kế hoạch 10%.

Yêu cầu:

- a) Tính tốc độ phát triển trung bình hàng năm về sản lượng?
- b) Hãy dự đoán sản lượng của doanh nghiệp trong 5 năm; biết rằng sản lượng năm gốc đạt 5.200 tấn sản phẩm?

Bài 5b:

Có tài liệu về doanh thu của công ty X như sau:

Cửa hàng	Thực tế năm 2014 so với thực tế năm 2013	Kế hoạch năm 2015 so với thực tế năm 2014	Thực tế năm 2015 so với kế hoạch năm 2015
A	110	115	104
B	105	120	102

Yêu cầu tính:

- a) Các tốc độ phát triển liên hoàn, định gốc và trung bình về doanh thu của cửa hàng trong thời gian từ 2013 - 2015?
- b) Giá trị tuyệt đối của 1% tăng qua các năm của từng cửa hàng. Biết rằng doanh thu thực tế năm 2013 của cửa hàng A là 2.000 triệu đồng; cửa hàng B là 3.000 triệu đồng?
- c) Tỷ lệ % hoàn thành kế hoạch năm 2015 tính cho toàn công ty?

Bài 6:

Có tài liệu về lượng hàng hoá tiêu thụ một loại hàng ở địa phương X như sau (đơn vị tính: tấn):

Tháng \ Năm	Năm	2013	2014	2015
1		1.200	1.300	1.240
2		1.180	1.250	1.200

3	1.500	1.600	1.450
4	1.800	1.900	1.780
5	2.700	2.600	2.630
6	3.400	3.300	3.000
7	4.400	4.500	4.300
8	5.000	4.900	4.800
9	4.000	3.800	3.900
10	2.100	2.050	1.800
11	1.500	1.400	1.450
12	1.000	1.100	1.070

- a) Hãy điều chỉnh các dãy số trên bằng phương pháp mở rộng khoảng cách thời gian (tháng -----> quý)?
- b) Tính các chỉ số thời vụ để nêu lên sự biến động về tình hình tiêu thụ loại hàng này?

Bài 7:

Có tài liệu của công ty X như sau:

Năm Chi tiêu	2009	2010	2011	2012	2013	2014	2015
Mức bán ra (tỷ đồng)	2000	2400	2600	2500	2980	3800	4200

Yêu cầu:

- a) Tính số trung bình di động cho từng nhóm 3 năm và lập thành dãy số mới?
- b) Trên cơ sở lượng tăng tuyệt đối trung bình và tốc độ phát triển trung bình hãy dự toán mức bán ra của công ty vào năm 2016 và 2017?
- c) Điều chỉnh dãy số bằng phương trình đường thẳng và dự đoán mức bán ra của công ty năm 2016 và 2017 trên cơ sở phương trình đường thẳng tính được?

Chương 7

CHỈ SỐ

Chỉ số là phương pháp được sử dụng rộng rãi trong nhiều ngành khoa học. Trong thống kê chỉ số là một phương pháp phân tích quan trọng và có những đặc thù riêng. Nội dung chương này sẽ trình bày cụ thể về các khái niệm, phân loại chỉ số, phương pháp tính các chỉ số cơ bản và giới thiệu một số chỉ tiêu kinh tế thông dụng có tính chất chỉ số được sử dụng phổ biến trong nền kinh tế.

7.1. Một số vấn đề chung về chỉ số

7.1.1. Khái niệm, đặc điểm và ý nghĩa của chỉ số

7.1.1.1. Khái niệm

Chỉ số trong thống kê là chỉ tiêu tương đối biểu hiện quan hệ so sánh giữa 2 mức độ của một hiện tượng nghiên cứu (hai thời gian, không gian khác nhau hoặc thực tế và kế hoạch).

Ví dụ:

+ Chỉ số về sản lượng của doanh nghiệp A năm 2015 so với năm 2014 là 120%, đây là quan hệ so sánh giữa giá trị sản lượng của doanh nghiệp trong 2 năm 2015 và 2014.

+ Chỉ số giá hàng hóa và dịch vụ tiêu dùng tháng 3/2014 so với tháng 2/2014 bằng 101% tức là tăng 1%.

+ Năm 2015, doanh nghiệp A đã hoàn thành kế hoạch về doanh thu 112%.

Hình thức biểu hiện của chỉ số là số tương đối (tức là so sánh cùng một loại chỉ tiêu) tuy nhiên không phải số tương đối nào cũng gọi là chỉ số, mà chỉ có 3 loại số tương đối sau mới là chỉ số: Số tương đối động thái, số tương đối kế hoạch, số tương đối so sánh (so sánh hiện tượng cùng loại ở thời gian khác nhau).

Như vậy, khái niệm chỉ số trong thống kê tương đối rộng, bao gồm các quan hệ so sánh để biểu hiện các mặt biến động của hiện tượng nghiên cứu. Trong thực tế, phương pháp chỉ số thường sử dụng để nghiên cứu sự biến động của các hiện tượng phức tạp, bao gồm các phần tử không thể cộng trực tiếp với nhau được.

7.1.1.2. Đặc điểm của phương pháp chỉ số

- Khi so sánh mức độ của hiện tượng giữa 2 thời gian hoặc không gian khác nhau của hiện tượng kinh tế - xã hội phức tạp, bao gồm các phần tử hay đơn vị không thể trực tiếp cộng được với nhau thì trước hết phải chuyển các phần tử hoặc đơn vị đó về dạng chung để có thể trực tiếp cộng được với nhau. Để thực hiện vấn đề này phải dựa trên cơ sở mối quan hệ giữa nhân tố cần nghiên cứu biến động của hiện tượng với các nhân tố khác có liên quan.

Ví dụ: Lượng hàng hoá tiêu thụ của mỗi mặt hàng vốn không thể cộng được với nhau vì các sản phẩm khác nhau có giá trị sử dụng và đơn vị tính khác nhau, nên phải chuyển về một dạng giống nhau để có thể cộng được, bằng cách nhân với nhân tố liên quan là giá cả (lượng hàng x giá) để được giá trị chung của sản phẩm.

- Để nghiên cứu biến động của một nhân tố, giả định các nhân tố khác không thay đổi tức là cố định các nhân tố đó ở cùng một thời kỳ, nhờ đó phương pháp chỉ số tạo khả năng loại trừ ảnh hưởng biến động của những nhân tố này để nêu lên sự biến động của riêng nhân tố cần nghiên cứu. Theo ví dụ ở trên, nếu tính chỉ số tổng hợp về lượng hàng hóa tiêu thụ thì trong chỉ số cho nhân tố lượng hàng hóa biến đổi còn giá cả thì được cố định ở cùng một thời kỳ nào đó; và nếu tính chỉ số tổng hợp về giá thì cho nhân tố giá biến đổi còn lượng hàng được cố định ở một thời kỳ nào đó. Trong chỉ số tổng hợp, nhân tố biến động gọi là nhân tố chỉ số hóa, nhân tố cố định gọi là quyền số của chỉ số.

7.1.1.3. Ý nghĩa của chỉ số

+ Biểu hiện sự biến động của hiện tượng qua thời gian bằng cách sử dụng chỉ số phát triển. Chỉ số phát triển được tính bằng cách so sánh mức độ của hiện tượng qua thời gian (kỳ nghiên cứu với kỳ gốc), như chỉ

số phát triển tính theo GDP; chỉ số phát triển sản xuất công nghiệp, sản xuất nông nghiệp, chỉ số phát triển về giá...

+ Biểu hiện sự so sánh của hiện tượng qua không gian khác nhau bằng cách sử dụng chỉ số không gian, như so sánh giá cả sinh hoạt của tỉnh A với tỉnh B.

+ Biểu hiện các nhiệm vụ kế hoạch và tình hình thực hiện kế hoạch, được vận dụng trong công tác xây dựng kế hoạch hoặc kiểm tra tình hình thực hiện kế hoạch thông qua chỉ số kế hoạch, so sánh năng suất lao động của Việt Nam với một số nước khác.

+ Phân tích mức độ ảnh hưởng và vai trò đóng góp của các nhân tố khác nhau đến sự biến động chung của hiện tượng phức tạp.

Ví dụ: Mức tiêu thụ hàng hoá = Giá cả x Lượng hàng hoá tiêu thụ. Như vậy, giá cả và lượng hàng hoá tiêu thụ là 2 nhân tố hình thành nên mức tiêu thụ hàng hoá. Khi giá cả tăng (giảm) thì mức tiêu thụ hàng hoá cũng tăng (giảm). Khi lượng hàng hoá tiêu thụ tăng (giảm) thì mức tiêu thụ hàng hoá cũng tăng (giảm). Vấn đề đặt ra ở đây là cần xác định xem khi giá cả tăng (giảm) thì làm cho mức tiêu thụ hàng hoá tăng (giảm) là bao nhiêu. Khi lượng hàng hoá tiêu thụ tăng (giảm) thì mức tiêu thụ hàng hoá tăng (giảm) bao nhiêu. Phương pháp chỉ số giúp ta đo lường được ảnh hưởng của từng nhân tố giá cả và lượng hàng tiêu thụ đến biến động chung của mức tiêu thụ hàng hoá cụ thể là bao nhiêu bằng hệ thống chỉ số được xây dựng trên cơ sở phương trình trên.

Như vậy, phương pháp chỉ số không những có khả năng tổng hợp từ phân tích sự biến động riêng biệt từng nhân tố của hiện tượng phức tạp, mà còn có thể xác định được vai trò đóng góp của từng nhân tố đó đến sự biến động chung của hiện tượng.

7.1.2. Phân loại chỉ số

Căn cứ theo các tiêu thức phân loại khác nhau, chỉ số được chia thành các loại sau đây:

- *Căn cứ vào phạm vi tính toán: Chỉ số được chia thành chỉ số đơn và chỉ số chung*

+ Chỉ số đơn (còn gọi là cá thể): Biểu hiện sự biến động của từng đơn vị, từng phần tử, hiện tượng cá biệt. Ví dụ: Chỉ số giá từng loại hàng hoá, chỉ số năng suất lao động từng công nhân, chỉ số khối lượng của từng sản phẩm...

Chỉ số đơn được sử dụng để nghiên cứu biến động từng phần tử hoặc từng đơn vị cấu thành nên tổng thể. Chỉ số đơn còn là cơ sở để tính chỉ số tổng hợp.

+ Chỉ số chung: Biểu hiện sự biến động các nhân tố của hiện tượng phức tạp bao gồm nhiều đơn vị, phần tử cá biệt. Ví dụ: Chỉ số chung về giá của một nhóm hàng lương thực - thực phẩm, hoặc tất cả các loại hàng hóa tiêu dùng trên thị trường, chỉ số khối lượng của tất cả các loại sản phẩm sản xuất ra trong năm.

- Căn cứ vào tính chất của chỉ tiêu nghiên cứu chỉ số được phân chia thành chỉ số chỉ tiêu chất lượng và chỉ số chỉ tiêu khối lượng:

+ Chỉ số chỉ tiêu chất lượng: Biểu hiện sự biến động của các chỉ tiêu chất lượng như: Giá cả, năng suất lao động, giá thành, tiền lương bình quân, GDP bình quân đầu người...

+ Chỉ số chỉ tiêu khối lượng: Biểu hiện sự biến động của các chỉ tiêu khối lượng như: Lượng hàng hoá tiêu thụ, số lượng công nhân, sản lượng, diện tích trồng lúa, lao động làm việc...

- Căn cứ hình thức biểu hiện của chỉ số, chia thành chỉ số ở dạng cơ bản và chỉ số ở dạng bình quân

+ Chỉ số ở dạng cơ bản: Là chỉ số được giữ nguyên công thức ban đầu khi xây dựng, không qua bất kỳ giai đoạn biến đổi nào khác (xem chỉ số tổng hợp về giá, chỉ số tổng hợp về lượng ở các phần sau).

+ Chỉ số ở dạng trung bình (viết gọn là chỉ số trung bình) được biến đổi từ dạng cơ bản về dạng như là số trung bình số học (trung bình cộng) gia quyền hay trung bình điều hòa gia quyền (chỉ số sản xuất công nghiệp là dạng chỉ số trung bình số học gia quyền được biến đổi từ chỉ số tổng hợp về khối lượng sản phẩm sản xuất công nghiệp).

- Căn cứ theo kỳ gốc so sánh, chia thành chỉ số liên hoàn và chỉ số định gốc:

+ Chỉ số liên hoàn: Là chỉ số tính cho nhiều thời kỳ liên tiếp nhau, trong đó mỗi chỉ số đều so sánh kỳ nghiên cứu (báo cáo) với thời kỳ liên kế trước đó.

+ Chỉ số định gốc: Là chỉ số tính cho nhiều thời kỳ khác nhau so với cùng một thời kỳ được chọn làm gốc.

7.2. Phương pháp tính chỉ số

Để minh họa phương pháp xây dựng và phân tích chỉ số trong thống kê, sau đây sẽ đề cập đến hai loại chỉ số thông dụng nhất của thống kê kinh tế - xã hội là chỉ số chỉ tiêu chất lượng với đại diện là chỉ số giá và chỉ số chỉ tiêu khối lượng với đại diện là chỉ số khối lượng sản phẩm, trên cơ sở sử dụng mối quan hệ: Mức tiêu thụ hàng hóa = giá cả x lượng hàng tiêu thụ làm ví dụ, sử dụng các ký hiệu như sau:

p: Giá đơn vị sản phẩm

q: Lượng hàng hoá tiêu thụ

i: Chỉ số đơn

I : Chỉ số tổng hợp

7.2.1. Phương pháp tính chỉ số phát triển

Chỉ số phát triển là số tương đối phản ánh quan hệ so sánh giữa hai mức độ của hiện tượng nghiên cứu ở hai thời gian khác nhau. Khi tính chỉ số phát triển là so sánh mức độ của hiện tượng theo thời gian, vì vậy sẽ ký hiệu “0” cho kỳ gốc (kỳ so sánh) và “1” cho kỳ nghiên cứu (kỳ nghiên cứu).

7.2.1.1. Phương pháp tính chỉ số đơn

Chỉ số đơn chỉ số phát triển thực chất là số tương đối động thái, do vậy chỉ số này được tính bằng cách so sánh mức độ kỳ nghiên cứu với mức độ kỳ gốc.

Công thức tính chỉ số đơn về giá: $i_p = \frac{p_1}{p_0}$

Công thức tính chỉ số đơn về chỉ tiêu khối lượng: $i_q = \frac{q_1}{q_0}$

Như vậy, chỉ số đơn được tính bằng cách so sánh mức độ của chỉ tiêu đó kỳ nghiên cứu với kỳ gốc. Ví dụ sau đây minh họa cho phương pháp tính chỉ số phát triển:

Bảng 7.2.1. Có tài liệu về tình hình tiêu thụ 3 mặt hàng ở công ty X như sau:

Loại hàng	Đơn vị tính	Giá bán lẻ (1000. Đ)		Lượng hàng hóa tiêu thụ	
		Kỳ gốc	Kỳ nghiên cứu	Kỳ gốc	Kỳ nghiên cứu
A	Kg	6	5	12000	15000
B	Mét	20	25	11400	12000
C	Lít	12	15	3500	4000

Từ số liệu trên, ta tính được các chỉ số đơn về giá của từng mặt hàng:

$$i_p(A) = \frac{5}{6} = 0,833 \text{ hoặc } 83,3\%; i_p(B) = \frac{25}{20} = 1,25 \text{ hoặc } 125\%$$

$$i_p(C) = \frac{15}{12} = 1,25 \text{ hoặc } 125\%$$

Như vậy, giá cả hàng hoá A kỳ báo cáo so với kỳ gốc bằng 83,3%, giảm 6,7%; Giá cả hàng hoá B kỳ nghiên cứu so với kỳ gốc bằng 125%, tăng lên 25%; Giá cả hàng hóa C kỳ nghiên cứu so với kỳ gốc bằng 125%, tăng 25%. Tương tự ta tính được các chỉ số đơn về lượng hàng tiêu thụ của từng mặt hàng:

$$i_q(A) = \frac{15000}{12000} = 1,25 \text{ hoặc } 125\%;$$

$$i_q(B) = \frac{12000}{11400} = 1,053 \text{ hoặc } 105,3\%;$$

$$i_q(C) = \frac{4000}{3500} = 1,142 \text{ hoặc } 114,2\%.$$

Như vậy, lượng tiêu thụ hàng hoá A kỳ nghiên cứu so với kỳ gốc bằng 125%, tăng 25%. Lượng tiêu thụ hàng hoá B kỳ nghiên cứu so với kỳ gốc bằng 105,3%, tăng 5,3% so với kỳ gốc. Lượng tiêu thụ hàng hoá C kỳ nghiên cứu so với kỳ gốc bằng 114,2, tăng 14,2%.

7.2.1.2. Phương pháp tính chỉ số tổng hợp

Chỉ số tổng hợp có dạng cơ bản và dạng trung bình.

7.2.1.2a. Chỉ số tổng hợp ở dạng cơ bản

Được dùng trong trường hợp có tài liệu về từng đơn vị trong tổng thể.

- Chỉ số tổng hợp về giá (I_p)

Với số liệu như ở bảng 7.2.1, ta thấy: Kỳ nghiên cứu so với kỳ gốc có giá cả mặt hàng B và C tăng lên, nhưng giá cả mặt hàng A giảm đi. Để tính chỉ số tổng hợp biểu hiện biến động chung về giá cả cho cả 3 mặt hàng (A, B, C), ta không thể cộng trực tiếp giá cả của 3 mặt hàng trên được với nhau vì số tổng cộng này không có ý nghĩa, và cách làm này không tính đến quyền số là khối lượng hàng hóa tiêu thụ, như vậy vai trò của từng mặt hàng (có lượng tiêu thụ khác nhau) ảnh hưởng đến chỉ số tổng hợp không được thể hiện. Do vậy, cần tiến hành xây dựng chỉ số tổng hợp về giá qua các bước như sau:

Bước 1: Chuyển từ một tổng thể bao gồm các phần tử không thể trực tiếp cộng được thành một tổng thể, trong đó các phần tử có thể cộng được bằng cách sử dụng nhân tố thông ước chung.

Cụ thể trong trường hợp tính chỉ số giá tổng hợp ta có: Giá cả x Lượng hàng hoá tiêu thụ = Mức tiêu thụ hàng hoá (Doanh thu). Lúc này mức tiêu thụ của từng loại hàng hoá đã có thể cộng được với nhau. Ta sẽ có được tổng mức tiêu thụ hàng hoá của cả 3 mặt hàng. Như vậy, nhân tố lượng hàng hoá tiêu thụ đóng vai trò là nhân tố thông ước chung, nhân tố thông ước chung đóng vai trò là quyền số của chỉ số tổng hợp.

Ta có: $p_0.q_0$: Là mức tiêu thụ hàng hoá của từng mặt hàng kỳ gốc.

$\Sigma p_0.q_0$: Tổng mức tiêu thụ hàng hoá kỳ gốc.

$p_1.q_1$: Mức tiêu thụ hàng hoá của từng mặt hàng kỳ nghiên cứu.

$\Sigma p_1.q_1$: Tổng mức tiêu thụ hàng hoá kỳ nghiên cứu.

Tổng mức tiêu thụ hàng hoá được tính chung cho nhiều loại hàng hoá khác nhau, có n mặt hàng thì tính được tổng mức tiêu thụ hàng hoá của n mặt hàng đó.

Nếu ta đem so sánh tổng mức tiêu thụ hàng hoá kỳ báo cáo với tổng mức tiêu thụ hàng hoá kỳ gốc ta được chỉ số tổng mức tiêu thụ hàng hoá.

$$I_{pq} = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_0} \quad (7.2.1)$$

Bước 2: Cố định lượng hàng tiêu thụ ở cả tử và mẫu số của công thức cùng theo một thời kỳ nào đó để nghiên cứu biến động riêng của nhân tố giá cả. Trong công thức (7.2.1), ta so sánh tổng mức tiêu thụ hàng hoá kỳ nghiên cứu so với kỳ gốc, chỉ số này phản ánh sự biến động của cả 2 nhân tố giá (p) và lượng (q). Mục tiêu ở đây là chỉ nghiên cứu ảnh hưởng biến động của riêng nhân tố giá, do vậy phải cố định nhân tố lượng hàng hóa tiêu thụ.

Từ công thức chỉ số tổng mức tiêu thụ hàng hoá, nếu ta cố định lượng hàng hoá tiêu thụ, ta sẽ tính được chỉ số tổng hợp về giá theo công thức:

$$I_p = \frac{\sum p_1 q}{\sum p_0 q} \quad (7.2.2)$$

Trong công thức (2) nhân tố (q) được gọi là quyền số của chỉ số giá tổng hợp. Trong công thức tính chỉ số, quyền số là đại lượng được cố định giống nhau ở tử số và mẫu số với hai chức năng cơ bản: Thứ nhất, làm cho các phần tử vốn không trực tiếp cộng được với nhau được chuyển về dạng đồng nhất để có thể tổng hợp được. Thứ hai, phản ánh vai trò hay tầm quan trọng của mỗi phần tử hay bộ phận trong toàn bộ tổng thể, tức là duy trì tỷ trọng của phần tử hay bộ phận đó tương xứng với vị trí của nó trong quá trình tổng hợp.

Từ công thức (7.2.2) nếu quyền số (q) cố định ở kỳ gốc, ta có công thức tính chỉ số giá như sau (chỉ số giá Laspeyres):

$$I_p = \frac{\sum p_1 \cdot q_0}{\sum p_0 \cdot q_0} \quad (7.2.3)$$

Trong công thức này, tử số là mức tiêu thụ hàng hóa kỳ gốc tính theo giá cả kỳ nghiên cứu; mẫu số là mức tiêu thụ hàng hóa kỳ gốc. Với số liệu ở bảng 7.2.1 ta tính toán được số liệu về mức tiêu thụ hàng hóa ở bảng 7.2.2.

Bảng 7.2.2: Tính toán tổng mức tiêu thụ hàng hóa kỳ gốc và kỳ nghiên cứu theo giá kỳ nghiên cứu

Loại hàng	Giá bán lẻ (1000 Đ)		Lượng hàng hóa tiêu thụ		Mức tiêu thụ hàng hoá	
	Kỳ gốc	Kỳ nghiên cứu	Kỳ gốc	Kỳ nghiên cứu	$p_0 \cdot q_0$	$p_1 \cdot q_0$
	(1)	(2)	(3)	(4)	(5=1x3)	(6=2x3)
A	6	5	12000	15000	72000	60000
B	20	25	11400	12000	228000	285000
C	12	15	3500	4000	42000	52500
					342000	397500

Theo số liệu bảng 7.2.2 và áp dụng công thức 7.2.3 ta tính được:

$$I_p = \frac{\sum p_1 \cdot q_0}{\sum p_0 \cdot q_0} = \frac{397500}{342000} = 1,16 \text{ hay } 116\%$$

Về tuyệt đối: $\sum p_1 \cdot q_1 - \sum p_0 \cdot q_1 = 397500 - 342000 = 55500$ đồng

Từ công thức (7.2.2) nếu quyền số (q) cố định ở kỳ nghiên cứu ta có công thức tính chỉ số giá (chỉ số giá Passche) như sau:

$$I_p = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1} \tag{7.2.4}$$

Trong đó:

- Tử số $\sum p_1 \cdot q_1$ là tổng mức tiêu thụ hàng hoá kỳ nghiên cứu.
- Mẫu số $\sum p_0 \cdot q_1$ là tổng mức tiêu thụ hàng hoá kỳ nghiên cứu tính theo giá kỳ gốc.

Hiệu số giữa tử số và mẫu số ($\sum p_1 \cdot q_1 - \sum p_0 \cdot q_1$) cho thấy với cùng một lượng hàng hoá bán ra như nhau, sự khác nhau giữa tử số và mẫu số

biểu hiện ảnh hưởng biến động của bản thân giá cả. Với số liệu ở bảng 7.2.1, ta tính toán được số liệu về mức tiêu thụ hàng hóa ở bảng 7.2.3.

Bảng 7.2.3: Tính tổng mức tiêu thụ hàng hóa kỳ nghiên cứu và kỳ nghiên cứu theo giá kỳ gốc

Loại hàng	Giá bán lẻ (1000 Đ)		Lượng hàng hóa tiêu thụ		Mức tiêu thụ hàng hóa	
	Kỳ gốc	Kỳ nghiên cứu	Kỳ gốc	Kỳ nghiên cứu	$p_1 \cdot q_1$	$p_0 \cdot q_1$
A	1	2	3	4	$5=2 \times 4$	$6=1 \times 4$
A	6	5	12000	15000	75000	90000
B	20	25	11400	12000	300000	240000
C	12	15	3500	4000	60000	48000
					435000	378000

Theo số liệu bảng 7.2.3 và áp dụng công thức 7.2.4 ta tính được:

$$I_p = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1} = \frac{435000}{378000} = 1,1508 \text{ hoặc } 115,08\%$$

Về tuyệt đối: $\sum p_1 \cdot q_1 - \sum p_0 \cdot q_1 = 435000 - 378000$

$$= 57000 \text{ (nghìn đồng)}$$

Như vậy, giá cả của 3 mặt hàng A, B, C kỳ nghiên cứu so với kỳ gốc tăng 15,05% làm cho tổng mức tiêu thụ hàng hoá của doanh nghiệp tăng 57.000 nghìn đồng. Hoặc giá cả của 3 mặt hàng tăng 15,08% so với kỳ gốc làm cho người tiêu dùng phải bỏ thêm một số tiền là 57.000 nghìn đồng.

Với kết quả tính toán trên cho thấy việc xác định thời kỳ quyền số khác nhau thì kết quả tính toán khác nhau, khi cơ cấu mặt hàng giữa hai thời kỳ thay đổi càng nhiều thì chênh lệch giữa hai cách tính trên càng lớn. Vì vậy, nhà thống kê học Fisher đề xuất sử dụng chỉ số tổng hợp về giá là mức độ trung bình nhân của hai chỉ số trên. Tuy nhiên, tính toán theo phương pháp của Fisher thì sẽ khó khăn hơn vì nguồn số liệu phải đầy đủ hơn (vừa phải tính $\sum p_0 q_1$ vừa phải tính $\sum p_1 q_0$).

- *Chỉ số tổng hợp chỉ tiêu khối lượng (I_q)*

Quá trình xây dựng công thức tính chỉ số chỉ tiêu khối lượng mà đại diện là chỉ tiêu lượng hàng hóa tiêu thụ tương tự như chỉ số tổng hợp chỉ tiêu giá.

Do lượng hàng hoá tiêu thụ có đơn vị tính khác nhau, nên không thể trực tiếp cộng được với nhau, do vậy phải lấy nhân tố giá làm thông ước chung để chuyển khối lượng hàng hoá tiêu thụ từ chỗ không thể cộng được với nhau thành mức tiêu thụ hàng hoá là đại lượng có thể cộng được, để có thể tổng hợp và so sánh.

Trong sự biến động của tổng mức tiêu thụ hàng hoá bao hàm ảnh hưởng của hai nhân tố (p) và (q). Mục đích nghiên cứu là xác định ảnh hưởng biến động của riêng nhân tố lượng tới mức tiêu thụ hàng hoá chung. Do vậy, nhân tố giá (p) phải được cố định cả ở tử số và mẫu số.

$$I_q = \frac{\sum p \cdot q_1}{\sum p \cdot q_0} \quad (7.2.5)$$

Trong đó, nhân tố giá tham gia với tư cách là nhân tố thông ước chung và cũng có vai trò là quyền số trong chỉ số lượng hàng hóa tiêu thụ. Quyền số này cũng phải được cố định ở cả tử số và mẫu số (kỳ gốc hoặc nghiên cứu).

Trường hợp quyền số (p) cố định ở kỳ gốc, ta có công thức tính chỉ số lượng hàng hóa tiêu thụ của Laspeyres (công thức 7.2.6).

$$I_q = \frac{\sum p_0 \cdot q_1}{\sum p_0 \cdot q_0} \quad (7.2.6)$$

Trong công thức này quyền số là giá ở kỳ gốc. Tử số $\sum p_0 q_1$ là tổng mức tiêu thụ hàng hoá kỳ nghiên cứu với giả định giá bán vẫn như kỳ gốc, mẫu số $\sum p_0 q_0$ là tổng mức tiêu thụ hàng hoá kỳ gốc. Với số liệu $\sum p_0 q_1$ ở bảng 7.2.3 và $\sum p_0 q_0$ ở bảng 7.2.2 ta tính được:

$$I_q = \frac{378000}{342000} = 1,1053 \text{ hoặc } 110,53\%$$

Tăng (giảm) tuyệt đối mức tiêu thụ hàng hóa:

$$\sum p_0 q_1 - \sum p_0 q_0 = 378.000 - 342.000 = 36.000 \text{ (nghìn đồng)}$$

Như vậy, lượng hàng hoá tiêu thụ chung của 3 mặt hàng của công ty kỳ nghiên cứu so với kỳ gốc tăng 10,53% làm cho tổng mức tiêu thụ hàng hoá tăng 36.000 nghìn đồng.

Trường hợp quyền số (p) cố định ở kỳ nghiên cứu, ta có công thức tính chỉ số lượng hàng hóa tiêu thụ của Paasche (công thức 7.2.7).

$$I_q = \frac{\sum p_1 \cdot q_1}{\sum p_1 \cdot q_0} \quad (7.2.7)$$

Cũng như chỉ số tổng hợp về giá, khi hai chỉ số tính với thời kỳ quyền số khác nhau có sự chênh lệch lớn, thì có thể vận dụng công thức Fisher, tức là trung bình nhân của hai chỉ số trên.

Tương tự, ta cũng có thể xây dựng công thức tính chỉ số tổng hợp cho chỉ tiêu khối lượng khác.

7.2.1.2b. Chỉ số tổng hợp ở dạng trung bình

Chỉ số tổng hợp ở dạng trung bình (viết gọn là chỉ số trung bình) là một dạng biến đổi của các chỉ số tổng hợp và công thức được trình bày dưới dạng là một số trung bình. Có hai loại chỉ số trung bình: Các chỉ số tổng hợp về giá có quyền số là lượng hàng tiêu thụ kỳ gốc và chỉ số tổng hợp về lượng có quyền số là giá cả kỳ gốc được biến đổi về dạng chỉ số trung bình số học gia quyền. Các chỉ số tổng hợp về giá có quyền số là lượng hàng tiêu thụ kỳ nghiên cứu và chỉ số tổng hợp về lượng có quyền số là giá cả kỳ nghiên cứu được biến đổi về dạng chỉ số trung bình điều hòa.

- *Chỉ số trung bình số học về giá có quyền số cố định ở thời kỳ gốc (từ công thức 7.2.3):*

$$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0} = \frac{\sum \frac{p_1}{p_0} p_0 q_0}{\sum p_0 q_0} = \frac{\sum i_p p_0 q_0}{\sum p_0 q_0} \quad (7.2.8)$$

$$\text{Và } \Delta_{pq}(p) = \left(\sum i_p p_0 q_0 - \sum p_0 q_0 \right)$$

Trong đó:

$i_p = \frac{p_1}{p_0}$: Chỉ số đơn về giá đóng vai trò lượng biến của chỉ số trung bình.

$p_0 q_0$: Giá trị hàng hóa tiêu thụ kỳ gốc đóng vai trò quyền số.

- *Chỉ số trung bình số học về lượng hàng hóa có quyền số cố định ở thời kỳ gốc (công thức 7.2.6):*

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0} = \frac{\sum \frac{q_1}{q_0} p_0 q_0}{\sum p_0 q_0} = \frac{\sum i_q p_0 q_0}{\sum p_0 q_0} \quad (7.2.9)$$

$$\text{Và } \Delta_{pq}(q) = \left(\sum i_q p_0 q_0 - \sum p_0 q_0 \right)$$

Trong đó:

$i_q = \frac{q_1}{q_0}$: Chỉ số cá thể về lượng hàng hóa tiêu thụ đóng vai trò lượng biến của số trung bình.

$p_0 q_0$: Giá trị hàng hóa tiêu thụ kỳ gốc đóng vai trò quyền số.

$\Delta_{pq}(p)$ và $\Delta_{pq}(q)$: Mức tiêu thụ hàng hóa tăng lên do tăng giá cả hoặc lượng hàng tiêu thụ.

Ví dụ: Từ số liệu về các chỉ số đơn về giá và lượng hàng tiêu thụ tính được ở trên và mức tiêu thụ hàng hóa của các loại mặt hàng ($p_0 q_0$) như số liệu bảng 7.2.2, áp dụng các công thức 7.2.8 và 7.2.9, ta tính được:

+ Chỉ số trung bình số học về giá (theo công thức 7.2.8):

$$I_p = \frac{(0,833 \times 72000) + (1,25 \times 22800) + (1,25 \times 42000)}{72000 + 22800 + 42000} = \frac{397500}{342000} = 1,1053$$

hoặc 110,53%

Và $\Delta_{pq}(p) = 397500 - 342000 = 55500$ nghìn đồng

+ Chỉ số trung bình số học về lượng hàng (theo công thức 7.2.9):

$$I_q = \frac{(1,25 \times 72000) + (1,053 \times 22800) + (1,142 \times 42000)}{72000 + 22800 + 42000} = \frac{378000}{342000} = 1,1026$$

hoặc 110,26%

Và $\Delta pq(q) = 378000 - 342000 = 36000$ nghìn đồng

Chú ý: Nếu ký hiệu $d_0 = \frac{p_0 q_0}{\sum p_0 q_0}$ (tỷ trọng mức tiêu thụ hàng hóa

của từng loại mặt hàng kỳ gốc) và thay vào các công thức (7.2.8) và (7.2.9) thì công thức chỉ số trung bình số học gia quyền về giá và lượng có dạng như sau:

$$I_p = \frac{\sum i_p d_0}{\sum d_0} \text{ và } I_q = \frac{\sum i_q d_0}{\sum d_0}$$

Trong đó: $\sum d_0 = 100\%$ (nếu tính bằng %), $\sum d_0 = 1$ (nếu tính bằng số lần).

Như vậy, các chỉ số tổng hợp về giá có quyền số là lượng hàng tiêu thụ kỳ nghiên cứu và chỉ số tổng hợp về lượng có quyền số là giá cả kỳ nghiên cứu được biến đổi về dạng chỉ số trung bình điều hòa gia quyền.

- *Chỉ số bình quân điều hòa về giá có quyền số cố định ở thời kỳ nghiên cứu (công thức 7.2.4):*

$$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1} = \frac{\sum p_1 q_1}{\sum \frac{p_0}{p_1} p_1 q_1} = \frac{\sum p_1 q_1}{\sum \frac{1}{i_p} p_1 q_1} \quad (7.2.10)$$

$$\text{Và } \Delta_{pq}(p) = \sum p_1 q_1 - \sum \frac{1}{i_p} p_1 q_1$$

Trong đó:

i_p : Chỉ số cá thể về giá đóng vai trò là lượng biến của mức độ trung bình.

$p_1 q_1$: Tổng giá trị hàng hóa tiêu thụ kỳ nghiên cứu đóng vai trò quyền số.

- Chỉ số bình quân điều hòa về lượng hàng hóa tiêu thụ có quyền số cố định ở thời kỳ nghiên cứu (công thức 7.2.7):

$$I_q = \frac{\sum p_1 q_1}{\sum p_1 q_0} = \frac{\sum p_1 q_1}{\sum \frac{q_0}{q_1} p_1 q_1} = \frac{\sum p_1 q_1}{\sum \frac{1}{i_q} p_1 q_1} \quad (7.2.11)$$

$$\text{Và } \Delta_{pq}(q) = \sum p_1 q_1 - \sum \frac{1}{i_q} p_1 q_1$$

Trong đó:

I_q : Chỉ số cá thể về giá đóng vai trò là lượng biến của mức độ trung bình.

$p_1 q_1$: Tổng giá trị hàng hóa tiêu thụ kỳ nghiên cứu đóng vai trò quyền số.

$\Delta_{pq}(p)$ và $\Delta_{pq}(q)$: Mức tiêu thụ hàng hóa tăng lên do tăng giá cả hoặc lượng hàng tiêu thụ.

Ví dụ: Cũng từ số liệu về các chỉ số đơn về giá cả và lượng hàng tiêu thụ tính được ở trên và mức tiêu thụ hàng hóa của các loại mặt hàng kỳ nghiên cứu ($\sum p_1 q_1$) như số liệu ở bảng 7.2.3, áp dụng các công thức (7.2.10) và (7.2.11) ta tính được:

+ Chỉ số trung bình điều hòa gia quyền về giá:

$$I_p = \frac{75000 + 300000 + 60000}{\frac{75000}{0,833} + \frac{300000}{1,25} + \frac{60000}{1,25}} = \frac{435000}{376000} = 1,1508 \text{ hoặc } 115,08\%$$

$$\text{Và } \Delta_{pq}(p) = 435000 - 378000 = 57000 \text{ nghìn đồng}$$

+ Chỉ số trung bình điều hòa gia quyền về lượng hàng tiêu thụ:

$$I_q = \frac{75000 + 300000 + 60000}{\frac{75000}{1,25} + \frac{300000}{1,053} + \frac{60000}{1,142}} = \frac{435000}{397500} = 1,0943 \text{ hoặc } 109,43\%$$

$$\text{Và } \Delta_{pq}(q) = 435000 - 397500 = 37500 \text{ nghìn đồng}$$

Các chỉ số trung bình được áp dụng trong các trường hợp có tài liệu về các chỉ số đơn và đặc biệt có ý nghĩa khi tiếp tục biến đổi quyền số của chỉ số về dạng “tỷ trọng giá trị của từng loại hàng hóa” để có thể sử dụng thuận lợi tỷ trọng đó khi tính toán và trong những trường hợp cần thiết có thể dùng tỷ trọng tương ứng để thay thế. Hiện nay, trong công tác Thống kê Công nghiệp ở Việt Nam đang tính chỉ số sản xuất Công nghiệp, đó chính là chỉ số ở dạng số trung bình số học gia quyền.

Trong trường hợp biết $d_1 = \frac{p_1 q_1}{\sum p_1 q_1}$ (tỷ trọng mức tiêu thụ hàng hóa kỳ nghiên cứu) thay vào công thức (7.2.10 và 7.2.11) thì công thức chỉ số trung bình điều hòa có dạng như sau:

$$I_p = \frac{\sum d_1}{\sum \frac{1}{i_p} d_1} \text{ và } I_q = \frac{\sum d_1}{\sum \frac{1}{i_q} d_1}$$

Chỉ số kế hoạch được tính toán tương tự như chỉ số phát triển.

7.2.2. Phương pháp tính chỉ số không gian

Chỉ số không gian biểu hiện quan hệ so sánh giữa hai mức độ của hiện tượng nghiên cứu ở hai không gian (địa phương) khác nhau. Về cơ bản phương pháp tính chỉ số không gian tương tự như đối với chỉ số phát triển. Điểm khác nhau cơ bản là vấn đề lựa chọn quyền số của chỉ số không gian. Cụ thể như sau:

7.2.2.1. Chỉ số đơn

- Chỉ số đơn về giá phản ánh quan hệ so sánh về giá của một mặt hàng ở hai không gian khác nhau.

Công thức tính:

$$i_{p(A/B)} = \frac{P_A}{P_B}$$

Trong đó: P_A : Là giá cả của thị trường A.

P_B : Là giá cả thị trường B.

- Chỉ số đơn về lượng hàng hóa tiêu thụ phản ánh quan hệ so sánh về lượng hàng hóa tiêu thụ ở hai không gian khác nhau.

Công thức tính:

$$i_{q(A/B)} = \frac{Q_A}{Q_B}$$

Trong đó: Q_A : Là lượng hàng hóa của thị trường A.

Q_B : Là lượng hàng hóa của thị trường B.

7.2.2.2. Chỉ số tổng hợp

- *Chỉ số tổng hợp về giá* phản ánh quan hệ so sánh của một nhóm hàng, hoặc tất cả các loại hàng hóa ở hai không gian khác nhau. Quyền số là chỉ tiêu khối lượng (q), không phải là lượng tiêu thụ thị trường A hay thị trường B, mà là tổng lượng hàng hoá tiêu thụ (cả thị trường A và thị trường B): $Q = q_A + q_B$

$$\text{Công thức tính: } I_p(A/B) = \frac{\sum p_A \cdot Q}{\sum p_B \cdot Q} = \frac{\sum p_A \cdot (q_A + q_B)}{\sum p_B \cdot (q_A + q_B)}$$

Trong đó, quyền số $Q = q_A + q_B$ là tổng khối lượng sản phẩm của từng mặt hàng ở cả hai thị trường.

- *Chỉ số tổng hợp về khối lượng sản phẩm* phản ánh quan hệ so sánh về khối lượng sản phẩm của một nhóm hàng hoặc tất cả các loại hàng hóa ở hai không gian khác nhau.

Trong thực tế có thể chọn giá cố định hoặc giá bình quân để làm quyền số. Nếu quyền số là giá bình quân, thì công thức tính chỉ số không gian chỉ số tổng hợp về khối lượng như sau:

$$I_q(A/B) = \frac{\sum \bar{p} \cdot q_A}{\sum \bar{p} \cdot q_B} \text{ trong đó: } \bar{p} = \frac{p_A \cdot q_A + p_B \cdot q_B}{q_A + q_B}$$

7.3. Hệ thống chỉ số

Hệ thống chỉ số là dãy các chỉ số có liên hệ với nhau, lập thành một đẳng thức nhất định, về trái là chỉ số chung phản ánh biến động của tất cả

các nhân tố còn về phải là các chỉ số nhân tố. Cơ sở của xây dựng hệ thống chỉ số là phương trình kinh tế biểu hiện mối liên hệ của các chỉ tiêu. Có nhiều loại hệ thống chỉ số, trong thực tế công tác thống kê thường sử dụng 3 loại hệ thống chỉ số sau: Hệ thống chỉ số tổng hợp, hệ thống chỉ số nghiên cứu biến động chỉ tiêu bình quân và hệ thống nghiên cứu chỉ tiêu tổng lượng biến.

7.3.1. Hệ thống chỉ số tổng hợp

Hệ thống chỉ số tổng hợp được xây dựng theo một số phương pháp khác nhau, trong đó phương pháp liên hoàn thường được sử dụng nhiều nhất.

7.3.1.1. Phương pháp liên hoàn

Hệ thống chỉ số này được hình thành dựa trên mối liên hệ thực tế giữa các chỉ tiêu và một số quy ước sau:

- Nhân tố chất lượng xếp trước, nhân tố khối lượng xếp sau theo thứ tự tính chất lượng giảm dần.

- Khi nghiên cứu ảnh hưởng biến động của một nhân tố thì cố định các nhân tố còn lại.

- Quyền số của nhân tố nghiên cứu là các nhân tố còn lại và lấy ở kỳ gốc đối với nhân tố xếp trước và kỳ nghiên cứu đối với nhân tố xếp sau.

- Trong hệ thống chỉ số tổng hợp nếu chỉ tiêu chung được cấu thành từ bao nhiêu nhân tố thì hệ thống chỉ số có bấy nhiêu chỉ số thành phần.

Ví dụ: Trên cơ sở mối liên hệ giữa mức tiêu thụ hàng hóa, giá cả và lượng hàng hóa tiêu thụ. Theo nguyên tắc của phương pháp liên hoàn, ta có hệ thống chỉ số sau:

+ Số tương đối:

$$\frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 \cdot q_1}{\sum p_0 \cdot q_1} \times \frac{\sum p_0 \cdot q_1}{\sum p_0 \cdot q_0} \quad (7.3.1.1)$$

$$I_{pq} = I_p \times I_q$$

+ Số tuyệt đối:

$$(\sum p_1q_1 - \sum p_0q_0) = (\sum p_1q_1 - \sum p_0q_1) + (\sum p_0q_1 - \sum p_0q_0)$$

$$\Delta_{pq} = \Delta_{pq}(p) + \Delta_{pq}(q)$$

Mối quan hệ giữa các số tương đối là quan hệ tích số; còn mối quan hệ giữa số tuyệt đối là quan hệ tổng.

Công thức (7.3.1.1) chính là một hệ thống chỉ số được xây dựng theo phương pháp liên hoàn, trong đó:

I_p, I_q : Là các chỉ số nhân tố;

I_{pq} : Là chỉ số chung (phản ánh biến động cả giá và lượng);

$\Delta_{pq}(p)$: Là chênh lệch tuyệt đối của mức tiêu thụ hàng hóa do ảnh hưởng của biến động giá cả;

$\Delta_{pq}(q)$: Là chênh lệch tuyệt đối của mức tiêu thụ hàng hóa do ảnh hưởng của tăng (giảm) lượng hàng hóa tiêu thụ;

Δ_{pq} : Là chênh lệch tuyệt đối của mức tiêu thụ hàng hóa nói chung.

Hệ thống chỉ số giúp ta phân tích vai trò ảnh hưởng của lượng hàng hóa tiêu thụ và giá cả đến sự biến động của mức tiêu thụ hàng hóa.

Ví dụ: Có số liệu về giá và lượng hàng hóa bán ra ở một cửa hàng và số liệu tính toán được trình bày ở bảng 7.3.1.1 như sau:

Bảng 7.3.1.1

Tên hàng	Giá bán (ng.đ)		Lượng tiêu thụ		Mức tiêu thụ hàng hóa		
	Kỳ gốc	Kỳ NC	Kỳ gốc	Kỳ nghiên cứu	p_0q_0	p_1q_0	p_1q_1
A	20	22	1000	1100	20.000	22.000	24.200
B	15	16,5	2000	2400	30.000	36.000	39.600
C	8	8	4000	6000	32.000	48.000	48.000
Σ					82.000	106.000	111.800

Thay số liệu đã tính toán ở bảng vào hệ thống chỉ số theo công thức 7.3.1.1 ta có:

$$\frac{111.800}{82.000} = \frac{111.800}{106.000} \times \frac{106.000}{82.000}$$

Giá trị các chỉ số: $1,3634 = 1,054 \times 1,2926$

Tăng (giảm) tuyệt đối mức tiêu thụ hàng hóa:

$$(111.800 - 82.000) = (111.800 - 106.000) + (106.000 - 82.000)$$

$$29.800 = 5800 + 24.000 \text{ (đơn vị 1000đ)}$$

Qua kết quả tính toán ở trên cho thấy:

Tổng mức tiêu thụ hàng hoá (doanh thu) kỳ nghiên cứu so với kỳ gốc của 3 mặt hàng tăng 36,34% tương ứng tăng 29,8 tr.đ là do 2 nguyên nhân:

+ Do giá cả chung của 3 mặt hàng kỳ nghiên cứu tăng lên 5,4% so với kỳ gốc làm cho doanh thu tăng lên 5,8 triệu đồng.

+ Lượng tiêu thụ của cả 3 mặt hàng tăng 29,26% nên làm cho doanh thu tăng 24 triệu đồng.

Như vậy, lượng hàng tăng lên là nhân tố chủ yếu làm tăng mức tiêu thụ hàng hóa.

7.3.1.2. Phương pháp biểu hiện biến động riêng biệt

Phương pháp này dựa trên cơ sở lý luận cho rằng tất cả các nhân tố đồng thời biến động và cùng liên hệ tác động lẫn nhau. Việc loại trừ ảnh hưởng biến động của các nhân tố không định nghiên cứu phải được tiến hành theo cùng một phương pháp, cùng một quan điểm. Phải làm thế nào để cho mỗi chỉ số nhân tố biểu hiện được ảnh hưởng biến động riêng biệt của nó đối với sự biến động của hiện tượng phức tạp, cho nên quyền số của các chỉ số này phải được chọn theo cùng một thời kỳ:

$$I_p = \frac{\sum p_1 q_0}{\sum p_0 q_0}$$

Chỉ số giá cả với quyền số kỳ gốc, biểu hiện ảnh hưởng biến động riêng biệt của nhân tố giá đối với biến động của mức tiêu thụ hàng hoá:

$$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}$$

Chỉ số lượng hàng hoá tiêu thụ biểu hiện ảnh hưởng biến động riêng biệt của nhân tố lượng hàng hoá tiêu thụ đối với sự biến động của mức tiêu thụ hàng hoá và quyền số là giá cả cùng cố định ở kỳ gốc.

Ta có hệ thống chỉ số xây dựng theo phương pháp biến động riêng biệt:

$$\frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_0 q_1}{\sum p_0 q_0} \times K \quad (7.3.1.2a)$$

$$I_{pq} = I_p \times I_q \times K$$

Trong đó: $K = \frac{\sum p_1 q_1 \cdot \sum p_0 q_0}{\sum p_1 q_0 \cdot \sum p_0 q_1}$ gọi là chỉ số liên hệ.

Thay giá trị của K vào 7.3.1.2a ta có:

$$\frac{\sum p_1 q_1}{\sum p_0 q_0} = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_0 q_1}{\sum p_0 q_0} \times \frac{\sum p_1 q_1 \cdot \sum p_0 q_0}{\sum p_1 q_0 \cdot \sum p_0 q_1} \quad (7.3.1.2b)$$

Tăng (giảm) tuyệt đối mức tiêu thụ:

$$\begin{aligned} (\sum p_1 q_1 - \sum p_0 q_0) &= (\sum p_1 q_0 - \sum p_0 q_0) + (\sum p_0 q_1 - \sum p_0 q_0) \\ &\quad + (\sum p_1 q_1 + \sum p_0 q_0 - \sum p_1 q_0 - \sum p_0 q_1) \end{aligned}$$

$$\Delta_{pq} = \Delta_{pq}(p) + \Delta_{pq}(q) + \Delta_{pq}(K)$$

Hai phương pháp trên nhằm giải quyết 2 nhiệm vụ nghiên cứu khác nhau, do vậy kết quả tính ra khác nhau. Hai phương pháp này không có gì mâu thuẫn nhau mà lại có thể bổ sung cho nhau làm cho phân tích thống kê càng sâu sắc hơn.

Trong thực tế phương pháp biến động liên hoàn được coi là phương pháp phân tích chủ yếu vì nó đơn giản và thích hợp với mục đích nghiên cứu hiệu quả kinh tế thực tế do biến động của các nhân tố.

7.3.2. Hệ thống chỉ số phân tích biến động của chỉ tiêu trung bình

Chỉ tiêu trung bình chịu ảnh hưởng biến động của 2 nhân tố: Bản thân tiêu thức bình quân hoá và kết cấu của tổng thể.

Ví dụ: Biến động tiền lương trung bình một công nhân trong doanh nghiệp là do biến động của bản thân mức lương của các loại công nhân và biến động kết cấu số công nhân có mức lương khác nhau; biến động giá thành trung bình 1 đơn vị sản phẩm là do tăng giảm bản thân giá thành của mỗi đơn vị sản phẩm và thay đổi tỷ trọng từng loại sản phẩm có lượng sản xuất khác nhau.

Việc đi sâu phân tích vai trò và ảnh hưởng biến động của các nhân tố đối với biến động của chỉ tiêu trung bình là nhiệm vụ cần thiết, nhằm chỉ ra cách thức để làm tăng (giảm) chỉ tiêu trung bình trong một tình huống phân tích cụ thể. Ví dụ: Để giảm giá hàng mua trung bình có thể bằng cách giảm đơn giá của từng lần mua hàng hoặc bố trí cơ cấu lượng hàng của từng lần mua hợp lý. Có thể thực hiện được nhiệm vụ phân tích này bằng phương pháp chỉ số. Cụ thể là dùng hệ thống chỉ số phân tích biến động chỉ tiêu trung bình. Biến động khi nghiên cứu mọi chỉ tiêu trung bình có 3 chỉ số hợp thành một hệ thống: Chỉ số cấu thành khả biến, chỉ số cấu thành cố định và chỉ số ảnh hưởng kết cấu.

(a) *Chỉ số cấu thành khả biến*: Được sử dụng để biểu hiện biến động của mức độ trung bình chung.

Nếu ký hiệu số trung bình kỳ báo cáo: $\bar{x}_1$ và số trung bình kỳ gốc: $\bar{x}_0$, ta có chỉ số trung bình chung ($I_{\bar{x}}$) được tính theo công thức:

$$I_{\bar{x}} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\frac{\sum x_1 f_1}{\sum f_1}}{\frac{\sum x_0 f_0}{\sum f_0}} \quad (7.3.2.1a)$$

Chỉ số trung bình chung phản ánh biến động của cả hai nhân tố: Lượng biến (x_i) và kết cấu tổng thể ($f_i / \sum f_i$).

(b) *Chỉ số cố định kết cấu (I_x)* được sử dụng để nghiên cứu ảnh hưởng biến động riêng của lượng biến (x_i). Để biểu hiện riêng ảnh hưởng của lượng biến, thì quyền số phải được cố định ở một kỳ. Nếu kết cấu quyền số được cố định ở kỳ báo cáo thì ta có công thức tính chỉ số cấu thành cố định (I_x):

$$I_x = \frac{\overline{x_1}}{\overline{x_{01}}} = \frac{\frac{\sum x_1 f_1}{\sum f_1}}{\frac{\sum x_0 f_1}{\sum f_1}} \tag{7.3.2.1b}$$

Chỉ số trung bình cố định kết cấu biểu hiện biến động của chỉ tiêu trung bình do ảnh hưởng riêng của lượng biến.

(c) *Chỉ số ảnh hưởng kết cấu:* Để nghiên cứu ảnh hưởng riêng biến động kết cấu đơn vị tổng thể:

$$I_s = \frac{\overline{x_{01}}}{\overline{x_0}} = \frac{\frac{\sum x_0 f_1}{\sum f_1}}{\frac{\sum x_0 f_0}{\sum f_0}} \tag{7.3.2.1c}$$

Chỉ số này biểu hiện biến động của chỉ tiêu trung bình do ảnh hưởng riêng của kết cấu (lượng biến cố định).

Ba chỉ số trên kết hợp với nhau lập thành một hệ thống chỉ số phân tích biến động chỉ tiêu trung bình do ảnh hưởng của 2 nhân tố: Bản thân lượng biến và kết cấu tổng thể:

$$I_x = L_x \times I_s \tag{7.3.2.1}$$

Cụ thể: $\frac{\overline{x_1}}{\overline{x_0}} = \frac{\overline{x_1}}{\overline{x_{01}}} \times \frac{\overline{x_{01}}}{\overline{x_0}}$

Tăng (giảm) tuyệt đối: $(\bar{x}_1 - \bar{x}_0) = (\bar{x}_1 - \bar{x}_{01}) + (\bar{x}_{01} - \bar{x}_0)$

$$\Delta_{\bar{x}} = \Delta_{\bar{x}}^{x_i} + \Delta_{\bar{x}}^{f_i/\sum f_i}$$

Trong đó: $\bar{x}_1 = \frac{\sum x_1 f_1}{\sum f_1}$, $\bar{x}_0 = \frac{\sum x_0 f_0}{\sum f_0}$, $\bar{x}_{01} = \frac{\sum x_0 f_1}{\sum f_1}$

Ví dụ: Có tài liệu về năng suất lao động và số công nhân của 2 phân xưởng sản xuất trong một doanh nghiệp như sau:

Bảng 7.3.2.1

Phân xưởng	Năng suất L.Đ (ng.đ)		Số công nhân (người)		$x_0 f_0$	$x_1 f_1$	$x_0 f_1$
	Kỳ gốc (x_0)	Kỳ ng.cứu (x_1)	Kỳ gốc (f_0)	Kỳ ng.cứu (f_1)			
A	1	2	3	4	5=1x3	6=2x4	7=1x4
Số 1	700	800	50	54	35.000	43.200	37.800
Số 2	750	780	50	52	37.500	40.560	39.000
Σ			100	106	72.500	83.760	76.800

Từ số liệu của bảng 7.3.2.1 ta tính các đại lượng:

$$\bar{x}_0 = \frac{\sum x_0 \cdot f_0}{\sum f_0} = \frac{72500}{100} = 725 \text{ (ng.đ)}$$

$$\bar{x}_1 = \frac{\sum x_1 \cdot f_1}{\sum f_1} = \frac{83760}{106} = 790,2 \text{ (ng.đ)}$$

$$\bar{x}_{01} = \frac{\sum x_0 \cdot f_1}{\sum f_1} = \frac{76800}{106} = 724,5 \text{ (ng.đ)}$$

Tiếp tục ta tính được:

- *Chỉ số trung bình chung* (phản ánh biến động của NSLĐ bình quân):

$$I_{\bar{x}} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{790,2}{725} = 1,089 \text{ lần hoặc } 108,9\%$$

Và $\Delta_{\bar{x}} = 790,2 - 725 = 65,2$ nghìn đồng

- *Chỉ số trung bình cố định kết cấu* (phản ánh biến động của bản thân năng suất lao động ảnh hưởng đến năng suất lao động trung bình):

$$I_{\bar{x}} = \frac{\bar{x}_1}{\bar{x}_{01}} = \frac{790,2}{724,5} = 1,091 \text{ lần hoặc } 109,1\%$$

Và $\Delta_{\bar{x}}(x) = 790,2 - 724,5 = 65,7$ nghìn đồng

- *Chỉ số trung bình ảnh hưởng kết cấu* (phản ánh biến động năng suất lao động bình quân do thay đổi kết cấu của công nhân):

$$I_s = \frac{\bar{x}_{01}}{\bar{x}_0} = \frac{724,5}{725} = 0,9993 \text{ lần (hay } 99,93\%).$$

Và $\Delta_{\bar{x}}(f / \Sigma f) = 724,5 - 725 = -0,5$ nghìn đồng.

Ba chỉ số trên lập thành một hệ thống chỉ số:

Số tương đối:

$$\frac{\bar{x}_1}{\bar{x}_0} = \frac{\bar{x}_1}{\bar{x}_{01}} \times \frac{\bar{x}_{01}}{\bar{x}_0}$$

$$\frac{790,2}{725} = \frac{790,2}{724,5} \times \frac{724,5}{725}$$

$$1,089 = 1,091 \times 0,9993$$

Số tuyệt đối:

$$(\bar{x}_1 - \bar{x}_0) = (\bar{x}_1 - \bar{x}_{01}) + (\bar{x}_{01} - \bar{x}_0)$$

$$(790,2 - 725) = (790,2 - 724,5) + (724,5 - 725)$$

$$65,2 = 65,7 - 0,5 \text{ (nghìn đồng).}$$

Từ kết quả tính toán ở trên ta có nhận xét: Năng suất lao động bình quân một công nhân toàn doanh nghiệp kỳ nghiên cứu so với kỳ gốc đã tăng lên 8,9% với mức độ tuyệt đối là 65,2 nghìn đồng/người. Trong đó:

+ Do bản thân năng suất lao động trung bình một công nhân của từng phân xưởng tăng lên làm cho năng suất lao động trung bình chung toàn doanh nghiệp tăng 9,1% tương ứng tăng 65,7 nghìn đồng/người.

+ Do sự thay đổi kết cấu số lượng công nhân từng phân xưởng kỳ nghiên cứu so với kỳ gốc làm cho năng suất lao động trung bình chung toàn doanh nghiệp giảm 0,07% tương ứng giảm 0,5 nghìn đồng/người.

Như vậy: Năng suất lao động trung bình chung toàn doanh nghiệp tăng lên chủ yếu là do bản thân năng suất lao động trung bình một công nhân trong từng phân xưởng tăng lên.

7.3.3. Hệ thống chỉ số nghiên cứu biến động của chỉ tiêu tổng lượng biến

Hệ thống chỉ số nghiên cứu biến động chỉ tiêu tổng lượng biến cho phép phân tích ảnh hưởng của 2 nhân tố: Chỉ tiêu trung bình ($\bar{x}$); và tổng số đơn vị của tổng thể ($\sum f_i$).

Ví dụ: Phân tích sự biến động của sản lượng do ảnh hưởng của năng suất lao động trung bình và số lượng công nhân hoặc phân tích biến động của tổng quỹ lương do ảnh hưởng biến động của tiền lương trung bình một công nhân và tăng (giảm) số lượng công nhân.

Ta có: $\sum x_i f_i = \bar{x} \times \sum f_i$

Trong đó:

$\bar{x}$: Là năng suất lao động bình quân.

$\sum f_i$: Là tổng số công nhân.

Trên cơ sở mối quan hệ này, ta có hệ thống chỉ số:

$$\frac{\sum x_1 \cdot f_1}{\sum x_0 \cdot f_0} = \frac{\bar{x}_1}{\bar{x}_0} \times \frac{\sum f_1}{\sum f_0}$$

$$I_{x.f} = I_{\bar{x}} \times I_{\sum f} \quad (7.3.3a)$$

$$\text{và } (\sum x_1 f_1 - \sum x_0 f_0) = (\bar{x}_1 - \bar{x}_0) \times \sum f_1 + (\sum f_1 - \sum f_0) \times \bar{x}_0$$

$$\Delta_{x.f} = \Delta_{\bar{x}.f}(\bar{x}) + \Delta_{x.f}(\sum f)$$

Trong đó: $I_{x.f}$: Chỉ số tổng lượng tiêu thức.

$I_{\bar{x}}$: Chỉ số chỉ tiêu trung bình.

$I_{\sum f}$: Chỉ số tổng tần số.

$\Delta_{x.f}$: Tổng lượng biến tăng lên.

$\Delta_{x.f}(\bar{x})$: Tổng lượng biến tăng do tăng chỉ tiêu trung bình.

$\Delta_{x.f}(\sum f)$: Tổng lượng biến tăng do tăng chỉ tiêu trung bình.

Với số liệu ở bảng (7.3.2.1) có thể tính toán và phân tích biến động của tổng sản lượng toàn doanh nghiệp kỳ báo cáo so với kỳ gốc bằng hệ thống chỉ số trên như sau:

Số tương đối: $I_{x.f} = I_{\bar{x}} \times I_{\Sigma f}$

$$\frac{83760}{72500} = \frac{790,2}{725} \times \frac{106}{100}$$

$$1,1553 = 1,089 \times 1,06$$

$$0,1553 = 0,089 + 0,06$$

Số tuyệt đối: $(83760 - 72500) = (790,2 - 725) \times 106 + (106 - 100) \times 725$

$$11260 = 6910 + 4350 \text{ (đơn vị 1000đ)}$$

Từ kết quả tính toán trên có thể thấy: Tổng sản lượng toàn doanh nghiệp kỳ nghiên cứu so với kỳ gốc tăng 15,53% tương ứng tăng 11,26 triệu đồng. Mức tăng này là do:

(1) Năng suất lao động trung bình một công nhân toàn doanh nghiệp kỳ nghiên cứu so với kỳ gốc tăng lên 8,9% tương ứng tăng 65,2 nghìn đồng một công nhân làm cho tổng sản lượng tăng 6,91 triệu đồng.

(2) Tổng số công nhân toàn doanh nghiệp kỳ nghiên cứu tăng so với kỳ gốc 6% tương ứng tăng 6 công nhân làm cho tổng sản lượng tăng 4,35 triệu đồng.

Như vậy, nhân tố chính làm cho tổng sản lượng tăng lên là do năng suất lao động trung bình 1 công nhân toàn doanh nghiệp tăng lên.

Khi kết hợp hệ thống chỉ số phân tích biến động chỉ tiêu bình quân đề cập ở mục 7.3.2 với hệ thống chỉ số theo công thức 7.3.3a, ta có hệ thống chỉ số phân tích biến động chỉ tiêu lượng biến với 3 nhân tố như sau:

Số tương đối:

$$\frac{\bar{x}_1 \Sigma f_1}{\bar{x}_0 \Sigma f_0} = \frac{\bar{x}_1 \Sigma f_1}{x_{01} \Sigma f_1} \times \frac{x_{01} \Sigma f_1}{x_{00} \Sigma f_1} \times \frac{\bar{x}_0 \Sigma f_1}{\bar{x}_0 \Sigma f_0} \quad (73.3b)$$

Hoặc: $I_{x.f} = I_{\bar{x}} \times I_{f/\Sigma f} \times I_{\Sigma f}$

Số tuyệt đối:

$$\begin{aligned} (\bar{x}_1 \Sigma f_1 - \bar{x}_0 \Sigma f_0) &= (\bar{x}_1 \Sigma f_1 \times \bar{x}_{01} \Sigma f_1 + (\bar{x}_{01} \Sigma f_1 - \bar{x}_0 \Sigma f_1) + (\bar{x}_0 \Sigma f_1 - \bar{x}_0 \Sigma f_0)) \\ \Delta_{x.f} &= \Delta_{x.f}(x) + \Delta_{x.f}(f/\Sigma f) + \Delta_{x.f}(\Sigma f) \end{aligned}$$

Từ số liệu đã có và tính toán ở mục 7.3.2, áp dụng công thức 7.3.3b ta tính được chỉ số phân tích chỉ tiêu tổng lượng biến:

+ Số tương đối:

$$\frac{790,2 \times 106}{725 \times 100} = \frac{790,2 \times 106}{724,5 \times 106} \times \frac{724,5 \times 106}{725 \times 106} \times \frac{725 \times 106}{725 \times 100}$$

$$\frac{83760}{72500} = \frac{83760}{76797} \times \frac{76797}{76850} \times \frac{76850}{72500}$$

$$1,1553 = 1,0907 \times 0,9993 \times 1,06$$

$$(+15,5\%) = (+9,1\%) + (-0,07\%) + (+6,0\%)$$

+ Số tuyệt đối:

$$(83760 - 72500) = (83760 - 76797) + (76797 - 76850) + (76850 - 72500)$$

$$11260 = 6963 + (-53) + 4350 \text{ (ng.đ.)}$$

Với kết quả tính toán ở trên ta thấy:

Tổng sản lượng toàn doanh nghiệp kỳ nghiên cứu so với kỳ gốc tăng 15,5% tương ứng tăng 11,26 triệu đồng là do 3 nguyên nhân:

+ Do bản thân năng suất lao động của công nhân từng phân xưởng tăng lên làm cho tổng sản lượng của toàn doanh nghiệp tăng 9,07% tương ứng với 6,963 triệu đồng, chiếm 61,84% ($6963/11260 \times 100$).

+ Do sự thay đổi kết cấu số lượng công nhân giữa các phân xưởng làm cho sản lượng chung toàn doanh nghiệp giảm 0,07% tương ứng giảm 0,53 triệu đồng, chiếm 4,71% ($-53/11260 \times 100$).

+ Do tổng số công nhân tăng 6% đã làm cho tổng sản lượng tăng 4,35 triệu đồng, chiếm 38,63% ($4530/11260 \times 100$).

7.4. Một số chỉ số thông dụng ở Việt Nam

7.4.1. Chỉ số giá tiêu dùng

Chỉ số giá tiêu dùng (CPI) là chỉ tiêu tương đối phản ánh sự biến động chung theo thời gian của mức giá hàng hóa và dịch vụ tiêu dùng cuối cùng của dân cư.

Giá tiêu dùng cuối cùng là số tiền mà dân cư phải trả khi mua một đơn vị hàng hóa hoặc dịch vụ phục vụ trực tiếp cho nhu cầu đời sống thường xuyên của mình. Giá tiêu dùng biểu hiện bằng giá bán lẻ hàng hóa trên thị trường và giá dịch vụ phục vụ sinh hoạt đời sống của dân cư. Trường hợp hàng hóa không có giá niêm yết thì giá tiêu dùng là giá thỏa thuận giữa người mua và người bán.

Giá tiêu dùng được thu thập ở từng địa phương trong toàn quốc theo danh mục quy định và tổng hợp cho toàn quốc phục vụ cho việc tính toán CPI.

Chỉ số giá tiêu dùng là một trong các chỉ tiêu quan trọng của hệ thống chỉ tiêu giá cả. CPI là cơ sở để đo mức lạm phát, thông qua đó để chính phủ đề ra các chính sách tài chính, tiền tệ và chính sách phân phối. CPI là chỉ tiêu quan trọng trong nghiên cứu thị trường, tác động đến những quyết định sản xuất kinh doanh của doanh nghiệp. CPI cũng là chỉ tiêu quan trọng cùng với các chỉ tiêu khác phản ánh mức sống dân cư.

Phương pháp tính CPI được tiến hành qua một số bước:

- *Bước 1:* Xác định kỳ gốc: Kỳ gốc là kỳ nào đó được chọn làm gốc so sánh, kỳ gốc có thể là gốc liên hoàn hoặc gốc cố định. Trong việc tính CPI thường sử dụng gốc liên hoàn, bởi người ta muốn biết nhịp độ tăng của giá cả một cách thường xuyên.

- *Bước 2:* Xác định “giỏ” hàng hóa và dịch vụ tiêu dùng. Giỏ hàng hóa này là danh mục các hàng hóa và dịch vụ tiêu dùng đại diện được quy định thống nhất.

- *Bước 3:* Xác định giá của từng mặt hàng, tức xác định (p) - kỳ gốc và kỳ nghiên cứu.

- *Bước 4:* Tính CPI theo công thức của Laspeyres:

$$CPI = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$$

Thông thường CPI được tính theo %.

Trong thực tế việc thu thập số liệu thống kê giá và lượng hàng hóa và dịch vụ tiêu dùng là rất phức tạp, đặc biệt là lượng (q). Chính vì vậy, phương án khả thi là điều tra chọn mẫu thông qua cuộc điều tra mức sống dân cư của cơ quan thống kê thực hiện và xác định cho từng nhóm mặt hàng một quyền số cố định, dựa trên tầm quan trọng (tỷ trọng) của nó trong tổng chi phí cho tiêu dùng cuối cùng của hộ dân cư.

Ở Việt Nam giai đoạn 2006 - 2010 xác định Danh mục hàng hóa và dịch vụ tiêu dùng được chia thành 4 cấp: Cấp I có 10 nhóm, cấp II có 34 nhóm, cấp III có 86 nhóm và cấp IV có 237 mặt hàng. Dựa theo danh mục này cơ quan thống kê các địa phương điều tra thu thập giá và tính CPI theo một cơ cấu quyền số cố định, cụ thể theo công thức sau:

$$CPI = \frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100 = \sum D_0 \frac{p_1}{p_0} \times 100$$

Trong đó: D_0 là quyền số cố định (theo kỳ gốc).

$$D_0 = \frac{p_0 q_0}{\sum p_0 q_0}$$

CPI ở Việt Nam được tính theo tháng và thường so với 3 tháng gốc (tháng trước, tháng cùng năm trước và tháng 12 năm trước) và được tính cho từng tỉnh, thành phố và cả nước có chia ra theo khu vực thành thị và nông thôn.

7.4.2. Chỉ số giá chứng khoán

Chỉ số giá chứng khoán là số tương đối phản ánh quan hệ so sánh giữa giá cổ phiếu bình quân kỳ nghiên cứu với giá bình quân kỳ gốc.

Chỉ số giá chứng khoán được coi là phong vũ biểu của nền kinh tế, nhất là đối với các thị trường chứng khoán lớn, là thông tin quan trọng đối với các nhà đầu tư và phân tích kinh tế. Các thị trường chứng khoán

đều xây dựng chỉ số giá chứng khoán cho riêng mình. Phương pháp xây dựng chỉ số giá chứng khoán cơ bản là dựa theo lý thuyết thống kê về chỉ số và số trung bình.

Chỉ số VN - Index là chỉ số giá cổ phiếu của sàn giao dịch chứng khoán Thành phố Hồ Chí Minh được tính căn bản theo phương pháp trung bình gia quyền với quyền số là khối lượng cổ phiếu niêm yết và tính giá bình quân kỳ gốc là 100 (100 điểm = 100%).

Cụ thể VN - Index được tính như sau:

$$VN.Index = \frac{\sum P_{ti} Q_{ti}}{\sum P_{oi} Q_{ti}} \times 100$$

Trong đó:

i: Là cổ phiếu niêm yết I tham gia tính chỉ số giá ($i = 1 \div n$);

n: Là số lượng cổ phiếu niêm yết đưa vào tính chỉ số giá;

P_{ti} : Giá của cổ phiếu I niêm yết tại thời điểm nghiên cứu (t);

P_{oi} : Giá của cổ phiếu I niêm yết tại kỳ gốc (0);

Q_{ti} : Khối lượng cổ phiếu I niêm yết tại thời điểm nghiên cứu (t).

7.4.3. Chỉ số sản xuất công nghiệp

Sản xuất công nghiệp là bộ phận trọng yếu của nền kinh tế quốc dân. Chính vì vậy, việc thống kê, tính toán, theo dõi sự biến động của chúng là rất cần thiết. Để làm việc này, ở nước ta cũng như trên thế giới cơ quan thống kê quốc gia thường xuyên tính Chỉ số sản xuất công nghiệp (IIP - Index of Industrial Production).

Chỉ số sản xuất công nghiệp được thực hiện theo quy trình như sau:

Chọn mẫu điều tra: Mẫu chọn đại diện cho cấp tỉnh, thành phố. Mỗi tỉnh, thành phố là một đơn mẫu được chọn theo các bước sau:

Bước 1: Chọn các ngành công nghiệp cấp 4 có quy mô giá trị tăng thêm lớn nhất từ cao xuống thấp, có giá trị tăng thêm cộng dồn chiếm từ 75% trở lên tổng giá trị tăng thêm của ngành công nghiệp cấp I để điều tra thường xuyên.

Bước 2: Trong các ngành công nghiệp cấp 4 được chọn mẫu ở bước 1, chọn các sản phẩm/mặt hàng công nghiệp lớn nhất từ cao xuống thấp, có giá trị tăng thêm cộng dồn chiếm từ 75% trở lên tổng giá trị tăng thêm của ngành công nghiệp cấp bốn để điều tra thường xuyên.

Bước 3: Trong các sản phẩm/mặt hàng công nghiệp được chọn mẫu ở bước 2, chọn các doanh nghiệp hoặc cơ sở sản xuất công nghiệp có sản lượng sản xuất lớn nhất từ cao xuống thấp đại diện (chiếm từ 75% trở lên tổng khối lượng sản phẩm/mặt hàng) để điều tra thường xuyên.

Công thức tính chỉ số sản xuất công nghiệp:

$$I_q = \frac{\sum i_q W_{q_0}}{\sum W_{q_0}}$$

Trong đó:

- I_q : Tốc độ phát triển sản xuất của một ngành, một khu vực hoặc một tỉnh, thành phố, tính theo khối lượng sản phẩm.
- i_q : Tốc độ phát triển của sản phẩm, hoặc của một ngành cấp dưới để tính cho ngành ở cấp cao hơn.
- W_q : Là quyền số được tính bằng giá trị tăng thêm.

- *Tính IIP cho cấp tỉnh, thành phố:*

Bước 1: Tính cho ngành cấp 4 mẫu:

$$I_{q_4} = \frac{\sum \frac{q_1}{q_0} W_{q_0}}{\sum W_{q_0}} = \frac{\sum i_q W_q}{\sum W_q}$$

Trong đó:

- I_{q_4} : Là tốc độ phát triển sản xuất của một ngành công nghiệp cấp 4.
- q_1 : Là khối lượng sản phẩm sản xuất của từng sản phẩm mẫu của ngành công nghiệp cấp 4 của kì báo cáo.
- q_0 : Giá trị tăng thêm của sản phẩm mẫu ở kỳ gốc.

Bước 2: Tính cho ngành cấp 1:

$$I_{q_1} = \frac{\sum I_{q_4} W_{q_{04}}}{\sum W_{q_{04}}}$$

Trong đó:

I_{q_1} : Là tốc độ phát triển sản xuất của một ngành công nghiệp cấp 1.

I_{q_4} : Là tốc độ phát triển sản xuất của một ngành công nghiệp cấp 4.

$W_{q_{04}}$: Là giá trị tăng thêm của ngành công nghiệp cấp 4 ở kỳ gốc.

Bước 3: Tính cho toàn ngành công nghiệp của một tỉnh, thành phố (hoặc cả nước):

$$I_q = \frac{\sum I_{q_1} W_{q_{01}}}{\sum W_{q_{01}}}$$

Trong đó:

I_q : Là tốc độ phát triển công nghiệp của tỉnh, thành phố (hoặc cả nước).

I_{q_1} : Là tốc độ phát triển sản xuất của 1 ngành công nghiệp cấp 1.

$W_{q_{01}}$: Là giá trị tăng thêm của ngành cấp 1 ở kỳ gốc.

Phương pháp tính chỉ số sản xuất công nghiệp là phương pháp đã và đang được các nước phát triển trên thế giới áp dụng rộng rãi và được xuất bản kết quả trong Niên giám Thống kê Liên hợp quốc.

TÓM TẮT CHƯƠNG 7

Trong công thức chỉ số tổng hợp biểu hiện biến động theo thời gian. Tùy theo điều kiện số liệu và mục đích nghiên cứu có thể sử dụng công thức tính chỉ số với thời kỳ quyền số khác nhau. Chỉ số tổng hợp Laspeyres chọn quyền số ở kỳ gốc. Chỉ số tổng hợp Pasch chọn quyền số ở kỳ nghiên cứu, chỉ số tổng hợp Fisher là trung bình nhân của 2 chỉ số tổng hợp Laspeyres và Passche. Chỉ số kế hoạch được tính toán tương tự

như chỉ số thời gian. Đối với chỉ số không gian, quyền số của chỉ số có những điểm khác biệt so với chỉ số thời gian.

Hệ thống chỉ số là một dãy các chỉ số có mối liên hệ với nhau, hợp thành một phương trình. Một hệ thống chỉ số gồm chỉ số chung và các chỉ số nhân tố, có 2 phương pháp xây dựng hệ thống chỉ số: phương pháp liên hoàn và phương pháp biến động riêng biệt. Hệ thống chỉ số được sử dụng để phân tích vai trò ảnh hưởng của từng nhân tố đối với sự biến động của chỉ tiêu tổng hợp. Trong thống kê thường sử dụng hệ thống chỉ số tổng hợp, hệ thống chỉ số trung bình và hệ thống chỉ số tổng lượng biến.

Trong nền kinh tế một số chỉ tiêu kinh tế quan trọng thể hiện bằng các chỉ số là: Chỉ số giá tiêu dùng, chỉ số giá chứng khoán, chỉ số sản xuất công nghiệp.

CÂU HỎI CHƯƠNG 7

1. Trình bày khái niệm, đặc điểm của chỉ số trong thống kê, phân biệt sự khác nhau giữa chỉ số và số tương đối?
2. Trình bày tác dụng của chỉ số, cho ví dụ minh họa?
3. Trình bày phương pháp xây dựng công thức tính chỉ số tổng hợp chỉ tiêu chất lượng và chỉ số tổng hợp chỉ tiêu khối lượng?
4. Trình bày công thức tính chỉ số tổng hợp bằng phương pháp trung bình. Trường hợp vận dụng cụ thể của các công thức đó?
5. Trình bày hệ thống chỉ số phân tích sự biến động của chỉ tiêu trung bình, ý nghĩa của hệ thống chỉ số đó. Cho ví dụ minh họa về trường hợp vận dụng cụ thể?
6. Trình bày hệ thống chỉ số phân tích sự biến động của tổng lượng tiêu thức, ý nghĩa của từng chỉ số trong hệ thống chỉ số?

BÀI TẬP CHƯƠNG 7

Bài 1:

Có tài liệu về tình hình sản xuất tại một xí nghiệp như sau:

Sản phẩm	Đơn vị tính	Giá thành đơn vị (1000đ)		Sản lượng	
		Kỳ gốc	Kỳ báo cáo	Kỳ gốc	Kỳ báo cáo
A	kg	32	30	4.000	4.200
B	cái	18	17,5	3.100	3.120
C	bộ	145	140	200	210

Yêu cầu tính:

- a) Chỉ số cá thể về giá thành và khối lượng sản phẩm?
- b) Chỉ số chung về giá thành và khối lượng sản phẩm? Số tiền tiết kiệm (hay chi phí thêm) của xí nghiệp do thay đổi giá thành và khối lượng sản phẩm sản xuất ra?
- c) Chỉ số chung về chi phí sản xuất của xí nghiệp? Phân tích nguyên nhân biến động của tổng chi phí sản xuất ở xí nghiệp trên? Cho nhận xét?

Bài 2:

Có tài liệu về giá cả và lượng hàng hoá tiêu thụ tại một thị trường như sau:

Hàng hoá	Đơn vị	Giá bán lẻ (1000đ)		Lượng hàng hoá tiêu thụ	
		Kỳ gốc	Kỳ báo cáo	Kỳ gốc	Kỳ báo cáo
A	kg	20	22	1.000	1.100
B	m	15	16,5	2.000	2.400
C	lít	8	8	4.000	6.000

Yêu cầu:

a) Tính chỉ số giá chung (theo công thức chỉ số liên hợp và chỉ số bình quân)?

b) Tính chỉ số giá chung về lượng hàng hoá tiêu thụ (theo công thức tính chỉ số liên hợp và chỉ số bình quân)?

c) Lập HTCS phân tích sự biến động của tổng mức tiêu thụ hàng hoá kỳ báo cáo so với kỳ gốc?

Bài 3:

Có tài liệu của công ty X như sau:

Tên hàng	Mức tiêu thụ hàng hoá kỳ báo cáo (trđ)	Tỷ lệ tăng (+) giảm (-) giá kỳ báo cáo so với kỳ gốc (%)
A	2.340	- 10
B	4.180	+ 10
C	4.900	- 2

Yêu cầu:

a) Tính chỉ số giá chung 3 loại hàng trên? Do giá cả thay đổi người tiêu dùng đã tiết kiệm hoặc vượt chi bao nhiêu khi so sánh kỳ báo cáo với kỳ gốc?

b) Phân tích nguyên nhân biến động của tổng mức tiêu thụ hàng hoá toàn công ty, biết rằng tổng mức tiêu thụ kỳ báo cáo tăng 20% so với kỳ gốc. Cho nhận xét?

Bài 4:

Có tài liệu sau đây:

Nhóm hàng	Tỷ trọng mức tiêu thụ hàng hoá kỳ báo cáo (%)	Tỷ lệ % tăng giá so với kỳ gốc
A	54	+ 8
B	14,8	- 4
C	10,2	+ 2
D	21,0	+ 5

Yêu cầu:

- a) Tính chỉ số giá chung của 4 nhóm hàng?
- b) Phân tích sự biến động của tổng mức tiêu thụ hàng hoá, biết rằng tổng mức tiêu thụ hàng hoá kỳ báo cáo là 9210 triệu đồng và tổng mức tiêu thụ hàng hoá kỳ báo cáo so với kỳ gốc tăng 20%?

Bài 5:

Có tài liệu của công ty X:

Tên hàng	Mức tiêu thụ hàng hoá kỳ gốc (trđ)	Tỷ lệ % tăng lượng hàng hoá tiêu thụ kỳ báo cáo so với kỳ gốc
A	1.700	+ 6
B	1.805	+ 8
C	2.500	+ 15

Yêu cầu tính:

- a) Chỉ số chung về lượng hàng hoá tiêu thụ?
- b) Chỉ số chung về giá cả, biết thêm rằng mức tiêu thụ hàng hoá chung kỳ báo cáo là 7050 (trđ). Phân tích sự biến động của tổng mức tiêu thụ hàng hoá?

Bài 6:

Một doanh nghiệp sản xuất 4 loại sản phẩm (A, B, C, D). Tổng chi phí sản xuất kỳ gốc là 2.400 trđ; trong đó sản phẩm A chiếm 20%, sản phẩm B chiếm 15%, sản phẩm C chiếm 35%, sản phẩm D chiếm 30%. Tổng chi phí sản xuất kỳ báo cáo là 2.640 trđ. Về khối lượng sản phẩm từng loại thì so với kỳ gốc, sản phẩm A tăng 16%, sản phẩm B tăng 20%, sản phẩm C tăng 10%, sản phẩm D tăng 25%.

Yêu cầu:

- a) Tính chỉ số chung về khối lượng sản phẩm?
- b) Tính chỉ số chung về giá thành?
- c) Phân tích nguyên nhân biến động của tổng chi phí sản xuất kỳ báo cáo so với kỳ gốc của doanh nghiệp?

Bài 7:

Có tài liệu về tình hình sản xuất một loại sản phẩm ở doanh nghiệp X như sau:

Phân xưởng	Kỳ gốc		Kỳ nghiên cứu	
	Sản lượng (cái)	Giá thành đơn vị (1000đ)	Sản lượng (cái)	Giá thành đơn vị (1000đ)
Số 1	2.000	100	6.000	95
Số 2	3.500	105	4.000	100
Số 3	4.500	110	2.000	105

Hãy dùng phương pháp chi số:

- Phân tích nguyên nhân biến động giá thành trung bình 1 sản phẩm chung của 3 phân xưởng?
- Phân tích nguyên nhân biến động của tổng chi phí sản xuất của toàn doanh nghiệp. Cho nhận xét?

Bài 8:

Có tài liệu của một doanh nghiệp thương mại X như sau:

Cửa hàng	Mức tiêu thụ hàng hoá (trđ)		Số lao động trung bình (người)	
	Kỳ gốc	Kỳ nghiên cứu	Kỳ gốc	Kỳ nghiên cứu
Số 1	1.755	2.160	15	18
Số 2	1.600	2.100	10	12
Số 3	1.500	2.475	15	15

Yêu cầu: Vận dụng phương pháp chi số:

- Phân tích các nhân tố ảnh hưởng đến sự biến động của NSLĐ trung bình toàn doanh nghiệp.
- Phân tích sự biến động tổng mức tiêu thụ hàng hoá của doanh nghiệp theo 3 nhân tố ảnh hưởng:
 - + NSLĐ?
 - + Kết cấu lao động?
 - + Số lượng lao động?

Chương 8

ĐIỀU TRA CHỌN MẪU

Trong điều tra thống kê, khi tổng thể thống kê quá lớn làm cho chi phí điều tra cao, tiêu tốn nhiều thời gian hoặc quy mô tổng thể không xác định thông thường phải sử dụng phương pháp điều tra chọn mẫu để thu thập tài liệu ban đầu. Có nghĩa là từ tổng thể chung ta chọn ra một tổng thể mẫu để tiến hành điều tra thực tế, kết quả trên mẫu điều tra được suy rộng cho tổng thể chung. Nội dung chương này sẽ trình bày một số vấn đề lý luận chung về điều tra chọn mẫu, phương pháp xác định phạm vi sai số chọn mẫu, suy rộng kết quả điều tra chọn mẫu, xác định số đơn vị tổng thể mẫu, các phương pháp tổ chức chọn mẫu ngẫu nhiên và quy trình tiến hành 1 cuộc điều tra chọn mẫu.

8.1. Khái niệm và ý nghĩa của điều tra chọn mẫu

Khi tiến hành điều tra thống kê, nếu xét theo phạm vi của đối tượng được điều tra thực tế, có thể phân loại điều tra làm hai loại là điều tra toàn bộ và điều tra không toàn bộ. Có nhiều loại điều tra không toàn bộ, trong đó điều tra chọn mẫu là loại điều tra không toàn bộ được sử dụng phổ biến nhất.

Điều tra chọn mẫu là một loại điều tra không toàn bộ, trong đó người ta chỉ chọn ra một số đơn vị trong toàn bộ các đơn vị của đối tượng nghiên cứu để điều tra thực tế. Các đơn vị này được chọn theo những quy tắc nhất định để đảm bảo tính đại biểu. Kết quả điều tra thường dùng để tính toán, suy rộng và đánh giá cho toàn bộ hiện tượng nghiên cứu.

Ví dụ: Để điều tra đánh giá chất lượng sản phẩm đồ hộp sản xuất của một doanh nghiệp, người ta chỉ tiến hành chọn ngẫu nhiên một số hộp trong lô sản phẩm đó để điều tra thực tế. Dựa vào kết quả điều tra được có thể tính toán, suy rộng và đánh giá chất lượng cho cả lô sản

phẩm sản xuất. Hay muốn tiến hành điều tra đánh giá mức sống dân cư của một địa phương, có thể chọn một số đủ lớn số hộ gia đình để thu thập tài liệu. Trên cơ sở tài liệu điều tra được ở những hộ gia đình đó suy rộng để đánh giá về mức sống dân cư của toàn địa phương...

Tuy chỉ điều tra một số đơn vị mà kết quả lại có thể suy rộng cho toàn bộ tổng thể. Về điều này, quy luật số lớn đã chỉ ra rằng: Nếu nghiên cứu một số tương đối lớn đơn vị cá biệt thì những biểu hiện ngẫu nhiên những đặc thù có tính đơn nhất sẽ bù trừ và triệt tiêu cho nhau, tính quy luật sẽ được biểu hiện rõ. Lý thuyết xác suất cũng chứng minh rằng: Sự sai khác giữa số bình quân của một số rất lớn các đại lượng ngẫu nhiên với kỳ vọng toán của nó là một lượng nhỏ tùy ý.

Trong điều tra chọn mẫu cần chọn một số đủ lớn các đơn vị để điều tra thực tế, có thể chọn theo hai cách: Chọn ngẫu nhiên và chọn phi ngẫu nhiên. Chọn ngẫu nhiên là việc chọn các đơn vị một cách khách quan, không phụ thuộc vào ý muốn chủ quan của người chọn, được gọi là *chọn mẫu ngẫu nhiên*. Chọn phi ngẫu nhiên là việc chọn các đơn vị không hoàn toàn khách quan, còn phụ thuộc vào ý muốn chủ quan của người chọn, được gọi là *chọn mẫu phi ngẫu nhiên*.

So với điều tra toàn bộ thì điều tra chọn mẫu có một số ưu điểm sau:

- Điều tra chọn mẫu thường nhanh hơn rất nhiều so với điều tra toàn bộ, vì điều tra ít đơn vị, nên các công việc chuẩn bị sẽ gọn, số lượng tài liệu ghi chép giảm đi, thời gian điều tra ghi chép, thời gian tổng hợp, phân tích sẽ được rút ngắn. Điều này làm cho điều tra chọn mẫu có tính kịp thời cao.

- Do chỉ điều tra thực tế ở một số đơn vị của hiện tượng nghiên cứu nên cần ít nhân viên điều tra và mọi chi phí sẽ giảm, do đó điều tra chọn mẫu tiết kiệm được khá nhiều sức người, vật tư và tiền của.

- Cũng do số đơn vị điều tra ít nên điều tra chọn mẫu cho phép mở rộng nội dung điều tra, đi sâu nghiên cứu nhiều mặt của hiện tượng nghiên cứu.

- Tài liệu thu thập được trong điều tra chọn mẫu sẽ có độ chính xác cao, bởi vì số nhân viên điều tra cần ít nên có thể chọn những người có

kinh nghiệm, có trình độ nghiệp vụ cao, đồng thời việc kiểm tra số liệu có thể tiến hành tỷ mỉ và tập trung, khiến cho các sai số do ghi chép giảm đi nhiều.

Điều tra chọn mẫu không đòi hỏi một tổ chức lớn như điều tra toàn bộ. Một cơ quan nhỏ cũng có thể tiến hành điều tra chọn mẫu.

Tóm lại, điều tra chọn mẫu có rất nhiều ưu điểm, sử dụng nó sẽ mang lại hiệu quả kinh tế cao. Song không vì thế mà cho rằng có thể dùng điều tra chọn mẫu để hoàn toàn thay thế điều tra toàn bộ và cũng không nên cho rằng kết quả suy rộng của điều tra chọn mẫu sẽ chính xác như kết quả điều tra toàn bộ. Do điều tra chọn mẫu chỉ tiến hành thu thập tài liệu ở một số đơn vị nên khi suy rộng cho toàn bộ hiện tượng nghiên cứu sẽ có những sai số so với kết quả điều tra toàn bộ. Đây chính là nhược điểm cơ bản của điều tra chọn mẫu. Vấn đề đặt ra là phải có những biện pháp nhằm giảm sai số đến mức thấp nhất, có thể chấp nhận được để tài liệu suy rộng phản ánh đúng thực tế hiện tượng nghiên cứu.

Điều tra chọn mẫu có thể được sử dụng linh hoạt trong nhiều trường hợp với mục đích khác nhau:

- Khi đối tượng nghiên cứu cho phép vừa có thể điều tra toàn bộ, vừa có thể điều tra chọn mẫu thì người ta thường tiến hành điều tra chọn mẫu để có kết quả nhanh và tiết kiệm chi phí hơn.

- Đối với những hiện tượng phức tạp có quy mô quá lớn, những hiện tượng khi điều tra có liên quan đến việc phá hủy đơn vị điều tra thì không thể điều tra toàn bộ mà phải sử dụng điều tra chọn mẫu.

- Kết hợp với điều tra toàn bộ để mở rộng nội dung điều tra và đánh giá kết quả điều tra toàn bộ. Trong một số cuộc điều tra (như tổng điều tra dân số) người ta đồng thời tổ chức điều tra chọn mẫu trong phạm vi nhỏ. Mục đích của cuộc điều tra chọn mẫu này là để mở rộng nội dung điều tra và kiểm tra kết quả của điều tra toàn bộ.

- Dùng để tổng hợp nhanh tài liệu điều tra toàn bộ. Trong các cuộc điều tra toàn bộ, nhất là trong tổng điều tra dân số thường phải tổng hợp rất nhiều chỉ tiêu khác nhau và qua nhiều cấp. Muốn hoàn thành khối

lượng như vậy phải mất nhiều thời gian, có khi hàng năm. Để có thông tin nhanh, phục vụ kịp thời cho yêu cầu của lãnh đạo, có thể dùng chọn mẫu để tổng hợp nhanh tài liệu.

- Điều tra chọn mẫu còn được sử dụng trong trường hợp muốn so sánh các hiện tượng với nhau mà chưa có tài liệu cụ thể, hoặc muốn kiểm định lại giả thiết đặt ra.

Điều tra chọn mẫu được ứng dụng rộng rãi trong nhiều ngành kinh tế quốc dân. Trong sản xuất công nghiệp có thể dùng điều tra chọn mẫu để kiểm nghiệm chất lượng sản phẩm, nghiên cứu tình hình năng suất lao động và tiền lương công nhân, tình hình sử dụng thời gian lao động, sử dụng thiết bị sản xuất và nguyên vật liệu... Trong nông nghiệp có thể dùng điều tra chọn mẫu để nghiên cứu tỷ suất nảy mầm của hạt giống điều tra năng suất và sản lượng cây trồng, điều tra chất lượng đàn gia súc... Trong thương mại dùng điều tra chọn mẫu để kiểm tra chất lượng hàng hóa, điều tra giá cả thị trường, nghiên cứu nhu cầu về hàng hóa tiêu dùng của nhân dân. Điều tra chọn mẫu còn được áp dụng trong nhiều ngành kinh tế quốc dân khác. Đặc biệt trong lĩnh vực xã hội, điều tra chọn mẫu càng được áp dụng nhiều hơn để nghiên cứu thái độ của mọi người đối với một vấn đề gì đó, điều tra tâm lý, xã hội...

8.2. Điều tra chọn mẫu ngẫu nhiên

8.2.1. Một số lý luận trong điều tra chọn mẫu ngẫu nhiên

8.2.1.1. Tổng thể chung và tổng thể mẫu

Tổng thể chung là tổng thể bao gồm toàn bộ các đơn vị thuộc đối tượng điều tra. Số đơn vị tổng thể chung thường được ký hiệu là N .

Tổng thể mẫu là tổng thể bao gồm một số đơn vị nhất định được chọn ra từ tổng thể chung để điều tra thực tế. Số đơn vị tổng thể mẫu thường được ký hiệu là n .

Ví dụ: Trong doanh nghiệp có 2.000 công nhân, người ta chọn ra 100 công nhân để điều tra về tình hình năng suất lao động. Như vậy, số đơn vị tổng thể chung $N = 2000$, số đơn vị tổng thể mẫu $n = 100$.

Tổng thể chung cũng như tổng thể mẫu đều có những tham số đặc trưng như: Số trung bình, tỷ lệ, phương sai...

Nội dung cơ bản của phương pháp chọn mẫu là dựa vào sự hiểu biết về tham số θ của tổng thể mẫu điều tra để suy rộng thành các tham số θ của tổng thể chung. Việc làm như vậy được gọi là ước lượng. Để có thể ước lượng hợp lý tối đa thống kê toán đã chứng minh rằng:

- Để ước lượng số trung bình chung $\bar{x}$, ta sử dụng số trung bình mẫu $\tilde{x}$ làm ước lượng.

- Để ước lượng tỷ lệ chung p ta sử dụng tỷ lệ mẫu w làm ước lượng.

- Để ước lượng phương sai chung σ^2 ta không dùng phương sai mẫu σ_0^2 mà sử dụng phương sai mẫu điều chỉnh S^2 làm ước lượng.

8.2.1.2. Chọn một lần và chọn nhiều lần

Việc chọn các đơn vị tổng thể mẫu có thể được thực hiện theo hai cách: Chọn một lần và chọn nhiều lần.

- Chọn một lần (còn gọi là chọn không hoàn lại hay không lặp lại) là khi mỗi đơn vị đã được chọn để đăng ký rồi sẽ được xếp riêng ra, không trả về tổng thể chung, do đó không có khả năng được chọn lại. Số đơn vị tổng thể chung sẽ giảm dần đi trong quá trình chọn từng đơn vị. Vì vậy, xác suất được chọn của mỗi đơn vị không bằng nhau, đồng thời việc chọn thứ tự từng đơn vị mẫu cho đủ n đơn vị cũng giống như lấy một lần n đơn vị ra để điều tra.

Gọi K là số khả năng thiết lập được tổng thể mẫu. Số khả năng đó trong chọn một lần được xác định theo công thức:

$$K \frac{N}{n} = \frac{N!}{n!(N-n)!}$$

- Chọn nhiều lần (còn gọi là hoàn lại hay chọn lặp lại) là khi mỗi đơn vị được chọn ra đăng ký rồi lại được trả về tổng thể chung. Như vậy, số đơn vị tổng thể chung không thay đổi trong suốt quá trình lựa chọn. Mỗi đơn vị đều có khả năng được chọn như nhau (xác suất được chọn của mỗi đơn vị bằng nhau). Đơn vị được chọn ra rồi vẫn có khả năng được chọn lại.

Trong chọn nhiều lần, số khả năng thiết lập tổng thể mẫu tính bằng công thức:

$$K = N^n$$

8.2.1.3. Sai số chọn mẫu và phạm vi sai số chọn mẫu

Bất kỳ cuộc điều tra nào cũng có sai số do ghi chép. Đó là sai số xảy ra do nhiều nguyên nhân: Do chưa hiểu đúng nội dung điều tra, do đo lường sai, do vô tình ghi chép sai v.v... Nếu sai số là ngẫu nhiên (khi nhiều hơn, khi ít hơn) thì nó sẽ bù trừ cho nhau khi điều tra một số lớn đơn vị, nhưng nếu có hệ thống (sai số về một phía) thì rất nguy hiểm vì càng điều tra nhiều đơn vị, sai số càng nhiều (thường xảy ra trong trường hợp dụng cụ đo lường sai). Sai số do ghi chép không thể xác định cụ thể nhưng có thể khắc phục bằng cách chuẩn bị tốt các khâu trước khi điều tra.

Trong điều tra chọn mẫu, ngoài sai số do ghi chép còn có sai số do tính chất đại biểu - sai số xảy ra do chỉ điều tra một số ít đơn vị mà kết quả lại suy rộng cho toàn bộ tổng thể.

Sai số do tính chất đại biểu cũng có thể là sai số có hệ thống hoặc sai số ngẫu nhiên. Sai số có hệ thống xảy ra do vi phạm nguyên tắc chọn, nghĩa là không đảm bảo tính khách quan khi chọn đơn vị điều tra. Điều này thường xảy ra đối với những trường hợp có dụng ý trước nên kết quả điều tra luôn lệch về một hướng nhất định (nhiều hơn hoặc ít hơn so với thực tế). Sai số này cũng không tính được nhưng có thể biết nó mang dấu (+) hoặc dấu (-).

Sai số ngẫu nhiên là sai số chỉ xuất hiện trong trường hợp các đơn vị của tổng thể mẫu được chọn theo quy tắc ngẫu nhiên. Do đó, khái niệm “*sai số chọn mẫu*” thường được hiểu là sai số chọn mẫu ngẫu nhiên. Đã gọi là ngẫu nhiên thì không thể biết trước sẽ lệch về phía nào. Sai số này sẽ tự điều chỉnh khi tăng số đơn vị điều tra.

Như vậy, mặc dù cuộc điều tra chọn mẫu được tổ chức hết sức khoa học, sai số ngẫu nhiên vẫn là điều không thể tránh khỏi. Tuy nhiên, trong thực tế người ta không để cho sai số này tự do phát sinh mà khống chế nó

trong phạm vi nhất định bằng cách tác động và các nhân tố ảnh hưởng đến sai số chọn mẫu.

Biện pháp quan trọng đầu tiên để giảm sai số chọn mẫu là tăng số đơn vị điều tra. Trong trường hợp các điều kiện khác nhau, số đơn vị điều tra càng nhiều, tính chất đại biểu của tổng thể mẫu càng cao và sai số chọn mẫu càng nhỏ. Với cùng một độ tin cậy, bình phương của sai số chọn mẫu tỷ lệ nghịch với số đơn vị mẫu điều tra. Nếu muốn giảm sai số chọn mẫu đi hai lần, phải tăng số đơn vị điều tra lên 4 lần. Nếu cứ tiếp tục tăng số đơn vị điều tra cho đến khi bằng số đơn vị tổng thể chung ($n=N$) thì sai số chọn mẫu không còn nữa vì điều tra chọn mẫu đã biến thành điều tra toàn bộ. Trong thực tế, số đơn vị mẫu không thể bằng số đơn vị chung vì như vậy sẽ mất ý nghĩa kinh tế của điều tra chọn mẫu. Do đó, phải căn cứ vào quy luật số lớn và yêu cầu của độ chính xác để xác định số đơn vị mẫu điều tra đủ khả năng khống chế sai số trong phạm vi cho phép.

Sai số chọn mẫu còn phụ thuộc vào tính đồng đều của các đơn vị tổng thể. Nếu trong tổng thể, lượng biến của các tiêu thức nghiên cứu chênh lệch nhau càng nhiều, sai số chọn mẫu càng lớn; nghĩa là sai số chọn mẫu tỷ lệ thuận với phương sai của tổng thể nghiên cứu. Trong trường hợp này muốn hạn chế sai số phải phân chia các đơn vị của hiện tượng nghiên cứu thành các tổ sao cho trong mỗi tổ các đơn vị đồng đều về chất, chênh lệch giữa các lượng biến ít đi, sau đó mới chọn các đơn vị điều tra trong tất cả các tổ.

Sử dụng phương pháp chọn mẫu thích hợp đối với từng hiện tượng nghiên cứu cũng là một biện pháp hạn chế sai số chọn mẫu. Muốn vậy, trước khi điều tra phải căn cứ vào mục đích nghiên cứu, phân tích tính chất, đặc điểm của hiện tượng để chọn phương pháp điều tra cho phù hợp.

Như vậy, sai số chọn mẫu là một trị số không cố định, phụ thuộc vào số đơn vị mẫu điều tra (n), vào trình độ đồng đều của tổng thể nghiên cứu (σ^2) và vào các phương pháp chọn mẫu khác nhau. Nếu ký hiệu sai số chọn mẫu là d thì ứng với mỗi mẫu cụ thể ta có một sai số:

$$d = \theta' - \theta$$

Vì ứng với mỗi tổng thể mẫu có một sai số nhất định nên cần phải tính sai số bình quân chọn mẫu (ký hiệu μ) để đánh giá độ chính xác của ước lượng, từ đó xây dựng khoảng tin cậy của ước lượng.

Thống kê toán đã chứng minh phương pháp xác định sai số trung bình chọn mẫu theo các công thức:

- Khi nhiệm vụ điều tra chọn mẫu là để suy rộng số trung bình về một tiêu thức nào đó, sai số trung bình chọn mẫu sẽ là:

+ Trường hợp chọn nhiều lần:

$$\mu_x = \sqrt{\frac{\sigma^2}{n}}$$

+ Trường hợp chọn một lần:

$$\mu_x = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$$

- Khi nhiệm vụ điều tra chọn mẫu là để suy rộng tỷ lệ theo một tiêu thức nào đó, sai số trung bình chọn mẫu sẽ là:

+ Trường hợp chọn nhiều lần:

$$\mu_p = \sqrt{\frac{p \cdot q}{n}} = \sqrt{\frac{w(1-p)}{n}}$$

+ Trường hợp chọn một lần:

$$\mu_p = \sqrt{\frac{p \cdot q}{n}} = \sqrt{\left(1 - \frac{n}{N}\right) \cdot p(1-p)}$$

Trong thực tế, khi tính sai số bình quân chọn mẫu thường không có tài liệu về phương sai chung nên phải thay thế bằng phương sai mẫu điều chỉnh S^2 .

Ví dụ: Tính sai số bình quân chọn mẫu: Trong một doanh nghiệp có 2000 công nhân, người ta chọn ra 100 công nhân để điều tra về năng suất lao động. Kết quả điều tra như sau:

Phân tổ công nhân theo NSLĐ (tấn)	Số công nhân (người) n_i	Trị số giữa x_i	$x_i n_i$	$x_i - \tilde{x}$	$(x_i - \tilde{x})^2$	$(x_i - \tilde{x})^2 n_i$
35-45	14	40	560	-18	324	4536
45-55	20	50	1000	-8	64	1280
55-65	42	60	2520	-2	4	168
65-75	20	70	1400	12	144	2880
75-85	4	80	320	22	484	1936
Cộng	100	-	5800	-		10800

Từ kết quả điều tra chọn mẫu trên, ta có:

- Năng suất lao động trung bình của công nhân được điều tra (số trung bình mẫu):

$$\tilde{x} = \frac{\sum x_i n_i}{\sum n_i} = \frac{5800}{100} = 58 \text{ tấn}$$

- Phương sai mẫu về năng suất lao động:

$$\sigma_0^2 = \frac{\sum (x_i - x)^2 n_i}{\sum n_i} = \frac{10800}{100} = 108$$

- Tỷ lệ số công nhân đạt năng suất lao động tiên tiến (từ 65 tấn trở lên) của mẫu điều tra (tỷ lệ mẫu):

$$w = \frac{m}{n} = \frac{20 + 4}{100} = 0,24 \text{ hay } 24\%$$

Vì phương sai mẫu điều chỉnh.

$$S^2 = \frac{n}{n-1} \sigma_0^2$$

và $w (1-w)' = \frac{n}{n-1} \cdot w (1-w)$

nên sai số trung bình chọn mẫu được xác định như sau:

Trường hợp chọn nhiều lần:

- Khi suy rộng về năng suất lao động trung bình:

$$\mu_x = \sqrt{\frac{\sigma_0^2}{n-1}} = \sqrt{\frac{108}{99}} = 1,044 \text{ tấn}$$

- Khi suy rộng tỷ lệ công nhân đạt năng suất lao động tiên tiến:

$$\mu_x = \sqrt{\frac{w(1-w)}{n-1}} = \sqrt{\frac{0,24 \cdot 0,76}{100-1}} = 0,0429 \text{ hay } 4,29\%$$

Trường hợp chọn một lần:

$$\mu_x = \sqrt{\frac{\sigma_0^2}{n-1} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{108}{99} \left(1 - \frac{100}{2000}\right)} = 0,036$$

$$\mu_x = \sqrt{\frac{w(1-w)}{n-1} \left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,24 \cdot 0,76}{99} \left(1 - \frac{100}{2000}\right)} = 0,0418 \text{ hay } 4,18\%$$

Các công thức sai số trung bình chọn mẫu trên đây biểu hiện trị số trung bình của các sai số chọn mẫu có thể gặp phải khi suy rộng tài liệu. Nhưng do tiến hành chọn ngẫu nhiên nên sai số này không phải là một trị số được xác định trước về dấu (+ hoặc -) mà phản ánh một phạm vi chênh lệch có thể nhiều hơn hoặc ít hơn so với các tham số của tổng thể chung. Như vậy, có nghĩa là chênh lệch giữa $\tilde{x}$ và $\bar{x}$; giữa w và p không phải hoàn toàn bằng μ mà nằm trong phạm vi $\pm\mu$.

Trong ví dụ về điều tra năng suất lao động ở trên đã tính được năng suất lao động trung bình mẫu $\tilde{x} = 58$ tấn, sai số trung bình chọn mẫu $\mu_x = 1,044$ tấn. Như vậy, có thể suy ra năng suất lao động trung bình chung của công nhân toàn doanh nghiệp là $\bar{x} = 58 \text{ tấn} \pm 0,044 \text{ tấn}$, nghĩa là trong phạm vi từ 56,956 tấn đến 59,044 tấn.

Theo chứng minh của toán học thì xác suất (hay trình độ tin cậy) của việc suy rộng nói trên chỉ mới bằng 0,6827, nghĩa là trong 10.000 trường hợp có 6827 trường hợp chắc chắn sai số chọn mẫu không vượt quá phạm vi $\pm\mu$.

Có thể nâng cao trình độ tin cậy của việc sai số chọn mẫu. Giả sử phạm vi này được mở rộng gấp 2 lần ($\pm 2\mu$) thì năng suất lao động trung bình của công nhân toàn doanh nghiệp suy ra sẽ là:

$$\bar{x} = 58 \pm 2 \cdot 1,044$$

$$\bar{x} = 58 \pm 2,088$$

Nghĩa là trong phạm vi từ 55,912 tấn đến 60,088 tấn. Trong trường hợp này theo chứng minh của toán thì trình độ tin cậy của việc suy rộng lên tới 0,9545.

Tương tự như vậy, nếu mở rộng phạm vi sai số chọn mẫu lên tới $\pm 3\mu$ thì trình độ tin cậy lên tới 0,9973.

Tóm lại, càng mở rộng phạm vi sai số chọn mẫu trình độ tin cậy của việc suy rộng càng tăng, đồng thời sai số chọn mẫu cũng tăng theo. Hệ số dùng để mở rộng đó được gọi là hệ số tin cậy (t), nó ứng với hàm xác suất $\phi(t)$ và đã được Lia-pu-nốp lập bảng tính sẵn.

Vậy có thể xác định phạm vi sai số chọn mẫu theo công thức:

$$\varepsilon = t\mu$$

Trong đó:

ε - Phạm vi sai số chọn mẫu.

t - Hệ số tin cậy.

Sau đây là một vài trị số tiêu biểu trong bảng tính sẵn theo hàm $\phi(t)$ của Lia-pu-nốp:

T	$\phi(t)$	t	$\phi(t)$	t	$\phi(t)$
1,0	0,6827	1,7	0,9109	2,4	0,9836
1,1	0,7287	1,8	0,9281	2,5	0,9876
1,2	0,7699	1,9	0,9426	2,6	0,9907
1,3	0,8064	2,0	0,9545	2,7	0,9931
1,4	0,8385	2,1	0,9643	2,8	0,9943
1,5	0,8684	2,2	0,9722	2,9	0,9963
1,6	0,8904	2,3	0,9766	3,0	0,9973

8.2.1.4. Xác định quy mô tổng thể mẫu

Trước khi tiến hành điều tra chọn mẫu phải xác định số đơn vị cần điều tra. Xác định số đơn vị mẫu điều tra phải đáp ứng hai yêu cầu:

- Bảo đảm sai số chọn mẫu nhỏ nhất.
- Làm cho chi phí thấp nhất.

Hai yêu cầu này đối lập nhau. Muốn bảo đảm tài liệu điều tra chính xác (sai số nhỏ nhất) phải có đủ kinh phí để điều tra một lượng đơn vị khá lớn. Ngược lại, với lượng kinh phí có hạn, chỉ điều tra được một số ít đơn vị thì phải chấp nhận một sai số nhất định.

Trong thực tế, người ta đã nghiên cứu nhiều cách xác định số đơn vị mẫu phù hợp với yêu cầu và điều kiện cụ thể của việc điều tra nghiên cứu. Thường căn cứ vào yêu cầu của độ chính xác phạm vi sai số chọn mẫu để tính số đơn vị mẫu cần điều tra.

Từ công thức tính phạm vi sai số chọn mẫu: $\varepsilon = t\sigma$ ta có:

- Số lượng đơn vị tổng thể mẫu khi suy rộng chỉ tiêu trung bình:

+ Trường hợp chọn nhiều lần:

$$n = \frac{t^2 \sigma^2}{\varepsilon_x^2}$$

+ Trường hợp chọn một lần:

$$n = \frac{N t^2 \sigma^2}{N \varepsilon_x^2 + t^2 \sigma^2}$$

- Số lượng đơn vị tổng thể mẫu khi suy rộng chỉ tiêu tỷ lệ:

+ Trường hợp chọn nhiều lần:

$$n = \frac{t^2 p \cdot q}{\varepsilon_p^2} \text{ với } q = 1 - p$$

+ Trường hợp chọn chọn một lần:

$$n = \frac{N t^2 p \cdot q}{N \varepsilon_p^2 + t^2 p \cdot q}$$

Từ các công thức trên đây có thể thấy được các nhân tố quyết định số đơn vị mẫu cần điều tra:

Thứ nhất: Số đơn vị mẫu (n) phụ thuộc vào phạm vi sai số chọn mẫu (ϵ). Nếu yêu cầu phạm vi sai số càng nhỏ thì số đơn vị mẫu cần điều tra càng nhiều và ngược lại. Theo công thức trên, số đơn vị mẫu tỷ lệ nghịch với bình phương phạm vi sai số chọn mẫu (ϵ^2).

Thứ hai: Số đơn vị mẫu (n) phụ thuộc vào hệ số tin cậy (t). Yêu cầu trình độ tin cậy của tài liệu suy rộng cao hay thấp thì hệ số tin cậy cũng phải lớn hay nhỏ một cách tương ứng và do đó số đơn vị mẫu cũng phải tăng hoặc giảm cho phù hợp. Trong công thức, số đơn vị mẫu tỷ lệ thuận với bình phương hệ số tin cậy (t^2).

Thứ ba: Số đơn vị mẫu (n) còn phụ thuộc vào tính chất đồng đều của hiện tượng nghiên cứu. Theo công thức ta thấy số đơn vị mẫu tỷ lệ thuận với phương sai của tổng thể chung (σ^2 và $p.q$). Nếu tiêu thức nghiên cứu biến thiên ít, phương sai tính ra nhỏ, số đơn vị mẫu cần điều tra sẽ giảm đi. Ngược lại, nếu tiêu thức biến thiên nhiều, phương sai sẽ lớn làm cho số đơn vị cần điều tra tăng lên.

Trong thực tế, khi tính số đơn vị mẫu cần điều tra thường không có tài liệu về phương sai chung, vì vậy việc áp dụng các công thức trên sẽ gặp khó khăn, có thể giải quyết bằng các cách sau:

- Lấy phương sai lớn nhất của các lần điều tra trước (nếu có).
- Lấy phương sai của các hiện tượng khác tương tự.
- Tổ chức điều tra thí điểm để tính phương sai.

Có thể ước lượng phương sai theo khoảng biến thiên (R) tùy theo đặc điểm phân phối của hiện tượng nghiên cứu... Nếu là phân phối chuẩn thì:

$$\sigma = \frac{X_{\max} - X_{\min}}{6}$$

Sau đây là một số ví dụ tính số đơn vị mẫu cần điều tra:

Ví dụ 1: Trong một xí nghiệp dệt gồm 4000 công nhân, người ta cần tính năng suất lao động trung bình trong một ngày bằng phương pháp chọn

mẫu, yêu cầu trình độ tin cậy là 0,997 và sai số không vượt quá 2 mét. Thực tế trong xí nghiệp cho thấy rằng, nhìn chung mỗi người ít nhất cũng dệt được 60 mét, người nhiều nhất dệt được không quá 90 mét. Vậy cần phải chọn ra ít nhất bao nhiêu công nhân để điều tra thực tế?

Ta thấy rằng số công nhân 4000 người là khá lớn, sự biến động về năng suất lao động coi như liên tục, có thể giả định rằng tổng thể chung được phân phối theo quy luật chuẩn và dùng công thức trên để ước lượng độ lệch tiêu chuẩn.

$$\sigma = \frac{90 - 60}{6} = 5 \text{ mét.}$$

Với yêu cầu trình độ tin cậy $\phi(t) = 0,997$, ta có hệ số tin cậy $t = 3$.

- Nếu chọn theo cách chọn nhiều lần, tính theo công thức:

$$n = \frac{t^2 \sigma^2}{\varepsilon^2} = \frac{3^2 \cdot 5^2}{2^2} = 57 \text{ công nhân}$$

- Nếu chọn theo cách chọn một lần, tính theo công thức:

$$n = \frac{N t^2 \sigma^2}{N \varepsilon^2 + t^2 \sigma^2} = \frac{4000 \cdot 3^2 \cdot 5^2}{4000 \cdot 2^2 + 3^2 \cdot 5^2} = 56 \text{ công nhân}$$

(nếu N khá lớn, thì tính n theo hai công thức cho ta kết quả gần bằng nhau).

Ví dụ 2: Một xí nghiệp đồ hộp tiến hành điều tra chọn mẫu để xác định tỷ lệ đồ hộp không đúng quy cách trong một đợt sản xuất. Yêu cầu xác suất của tài liệu suy rộng là 0,9545, phạm vi sai số không vượt quá 4%. Trong ba lần điều tra trước tỷ lệ đồ hộp không đạt tiêu chuẩn là 3%, 4% và 5%. Hãy tính số đồ hộp cần điều tra lần này.

Để có phương sai lớn nhất ta sử dụng tỷ lệ $p = 0,05$. Với yêu cầu độ tin cậy 95,45% tra bảng phân phối chuẩn ta được $t = 2$. Vì số đồ hộp sản xuất ra thường rất nhiều (N lớn) nên dù có chọn theo phương pháp chọn một lần vẫn có thể sử dụng công thức:

$$n = \frac{2^2 \times 0,05 \times 0,95}{0,04^2} = 119 \text{ hộp}$$

Kiểm tra tính chất đại biểu của tổng thể mẫu:

Sau khi xác định và lựa chọn được số đơn vị mẫu cần điều tra, người ta thường kiểm tra tính chất đại biểu của tổng thể đó. Đây là việc làm cần thiết vì nếu tổng thể mẫu không đại biểu cho tổng thể chung thì kết quả suy rộng sẽ không chính xác.

Việc kiểm tra tính chất đại biểu của tổng thể mẫu có thể tiến hành bằng nhiều cách:

- So sánh một số chỉ tiêu bình quân hoặc chỉ tiêu tương đối của tổng thể mẫu được chọn với các chỉ tiêu tương ứng của tổng thể chung ($\tilde{x}$ với $\bar{x}$; W với P). Đương nhiên là giữa các chỉ tiêu này thường có chênh lệch, nhưng nếu mức độ chênh lệch không vượt quá phạm vi cho phép (thường là $\pm 5\%$) thì tổng thể mẫu được coi như đủ tính chất đại biểu. Nếu mức độ chênh lệch vượt quá phạm vi $\pm 5\%$, tổng thể mẫu chưa đủ tính chất đại biểu, do đó phải thay thế một số đơn vị hoặc tăng thêm số đơn vị điều tra.

- Kiểm tra và xử lý những số liệu nghi ngờ. Nếu trong quá trình chọn các đơn vị điều tra gặp phải một vài đơn vị có số liệu quá lớn hay quá bé, ta cần kiểm tra và xử lý như sau:

- + Nếu số đơn vị có số liệu đột xuất đó không nhiều (trường hợp cá biệt) thì nên loại bỏ để thay thế bằng đơn vị khác.

- + Nếu số đơn vị có số liệu đột xuất không phải là cá biệt (gặp một số lần) thì nên chọn phân loại để đảm bảo có một số đơn vị mẫu nhất định đại diện cho loại đó trong tổng thể chung.

Ngoài ra người ta còn có thể so sánh sự phân phối các đơn vị theo tiêu thức nghiên cứu của tổng thể chung và tổng thể mẫu. Cách so sánh này phức tạp song bảo đảm tính đại biểu hơn.

8.2.1.5. Suy rộng kết quả điều tra chọn mẫu:

Suy rộng kết quả điều tra chọn mẫu là tính toán các tham số của tổng thể chung trên cơ sở tài liệu thu thập được trong điều tra chọn mẫu. Có hai phương pháp suy rộng sau đây:

- *Phương pháp tính đối trực tiếp*: Phương pháp này được áp dụng khi người ta dùng các số bình quân hay số tương đối của tổng thể mẫu để tính ra các tham số tương ứng của tổng thể chung. Cách tính như sau:

$$\bar{x} = \tilde{x} \pm \varepsilon_x \text{ hay } \tilde{x} - \varepsilon_x \leq \bar{x} \leq \tilde{x} + \varepsilon_x$$

$$p = w \pm \varepsilon_p \text{ hay } w - \varepsilon_p \leq p \leq w + \varepsilon_p$$

Ví dụ: Giả sử sau khi điều tra chọn mẫu số đồ hộp (119 hộp) tính được tỷ lệ đồ hộp không đúng quy cách là 10%. Với xác suất 0,9545 tính ra phạm vi sai số chọn mẫu là $\varepsilon_p = 0,02$. Như vậy, suy ra tỷ lệ đồ hộp không đúng quy cách của cả đợt sản xuất của xí nghiệp sẽ nằm trong phạm vi:

$$w - \varepsilon_p \leq p \leq w + \varepsilon_p$$

$$0,10 - 0,02 \leq p \leq 0,1 + 0,02$$

$$0,08 \leq p \leq 0,12 \text{ hay } 8\% \leq p \leq 12\%$$

Vậy, tỷ lệ sản phẩm đồ hộp không đúng quy cách của cả đợt sản xuất nằm trong khoảng từ 8% đến 12%.

- *Phương pháp hệ số điều chỉnh*: Phương pháp này thường được dùng để xác minh kết quả điều tra toàn bộ. Căn cứ vào kết quả của điều tra toàn bộ và điều tra chọn mẫu, tính ra tỷ lệ chênh lệch rồi dùng tỷ lệ này làm hệ số điều chỉnh kết quả điều tra toàn bộ.

Ví dụ: Trong Tổng điều tra dân số 1/4/2009 người ta đồng thời tiến hành điều tra chọn mẫu 5% số địa bàn điều tra với mục đích nghiên cứu thêm một số tiêu thức, đồng thời kiểm tra kết quả tổng điều tra dân số. Giả sử, theo kết quả của Tổng điều tra dân số thì tổng số dân của tỉnh X vào thời điểm 1/4/2009 là 1765200 người, trong đó chỉ tính riêng 5% địa bàn điều tra (là những địa bàn chọn làm mẫu) là 95.260 người. Trong khi đó kết quả điều tra chọn mẫu trên 5% số địa bàn toàn bộ đã bỏ sót mất 48 người trong các địa bàn đó.

$$\text{Tỷ lệ bỏ sót là: } \frac{48}{96260} = 0,0005 \text{ hay } 0,05\%$$

Có thể dùng tỷ lệ này làm hệ số điều chỉnh lại kết quả điều tra toàn bộ. Như vậy, trong toàn tỉnh số người bị bỏ sót có thể lên tới:

$$0,0005 \times 1.765.200 = 883 \text{ người.}$$

Vậy, số dân thực tế của tỉnh vào thời điểm 1/4/2009 là:

$$1.765.200 + 883 = 1.766.083 \text{ người.}$$

8.2.2. Các phương pháp tổ chức chọn mẫu ngẫu nhiên thường dùng trong thống kê

Chọn các đơn vị mẫu từ tổng thể chung có thể tiến hành theo nhiều cách khác nhau. Hệ thống tổ chức chọn các đơn vị mẫu từ tổng thể chung gọi là phương pháp tổ chức chọn mẫu. Thống kê thường sử dụng các phương pháp tổ chức chọn mẫu sau:

- Chọn ngẫu nhiên đơn thuần.
- Chọn máy móc.
- Chọn phân loại.
- Chọn cả khối.
- Chọn phân tầng.

Mỗi phương pháp tổ chức chọn mẫu có đặc điểm riêng và cách tính sai số chọn mẫu riêng.

8.2.2.1. Phương pháp chọn ngẫu nhiên đơn thuần

Phương pháp chọn ngẫu nhiên đơn thuần là phương pháp tổ chức chọn các đơn vị mẫu trong tổng thể chung một cách hết sức ngẫu nhiên không qua một sự sắp xếp nào cả. Người ta lần lượt chọn ra từng đơn vị bằng cách rút thăm, quay số hoặc chọn theo bảng số ngẫu nhiên. Mỗi đơn vị tổng thể chung có thể được chọn một lần (không lặp lại) hoặc chọn nhiều lần (lặp lại).

Khi tính sai số bình quân chọn mẫu có thể dùng các công thức đã được trình bày ở phần trên.

Phương pháp chọn ngẫu nhiên đơn thuần là phương pháp đơn giản nhất nhưng nó chỉ cho ta kết quả tốt khi tổng thể nghiên cứu tương đối

đồng đều. Nếu tổng thể nghiên cứu có kết cấu phức tạp thì không nên dùng phương pháp này vì chọn như vậy số đơn vị tổng thể mẫu không được phân phối đều trong tổng thể chung, tính chất đại biểu của tổng thể mẫu sẽ không cao. Mặt khác nếu tổng thể chung có quy mô lớn thì việc đặt số hiệu thăm cho tất cả các đơn vị cũng gặp khó khăn. Chính vì những nhược điểm trên mà phương pháp chọn ngẫu nhiên đơn thuần ít được áp dụng để điều tra các hiện tượng kinh tế - xã hội.

8.2.2.2. *Phương pháp chọn máy móc (chọn hệ thống)*

Phương pháp chọn máy móc là phương pháp tổ chức chọn mẫu trong đó mỗi đơn vị được chọn căn cứ vào từng khoảng cách nhất định. Trước hết, người ta sắp xếp các đơn vị tổng thể chung theo một thứ tự nào đó như sắp xếp theo vần A, B, C... của tên gọi, theo thứ tự địa danh, theo quy mô từ nhỏ đến lớn v.v... Sau đó, lần lượt chọn các đơn vị theo thứ tự một cách máy móc, tức là cứ sau mỗi khoảng cách nhất định lại chọn ra một đơn vị. Khoảng cách này được xác định bằng cách lấy số đơn vị tổng thể chung chia cho số đơn vị tổng thể mẫu.

$$d = \frac{N}{n}$$

Giả sử cần chọn 100 công nhân trong số 2000 công nhân của doanh nghiệp để điều tra về năng suất lao động. Có thể sắp xếp 2000 công nhân này thành một danh sách theo vần của tên gọi (A, B, C...). Khoảng cách để chọn ra từng người là $2000 : 100 = 20$. Nghĩa là cứ cách 20 người trong danh sách lại chọn ra một người. Người công nhân đầu tiên được chọn ra trong số 20 người đầu của danh sách bằng cách chọn ngẫu nhiên đơn thuần. Giả sử trong 20 người này rút thăm được người thứ 12, thì những người chọn tiếp sau sẽ là người thứ 32, 52, 72...

Trong thực tế có những hiện tượng mà các đơn vị của nó đã được sắp xếp sẵn và trong thời gian điều tra vẫn không thay đổi vị trí. Với những hiện tượng như vậy có thể lợi dụng sự bố trí sẵn có để chọn theo khoảng cách đã tính.

Ví dụ: Khi điều tra chọn mẫu dân số, cứ theo thứ tự số nhà trên các đường phố mà chọn theo khoảng cách, khi điều tra năng suất lúa, sau khi

xác định khoảng cách chọn, tiến hành cắm điểm trên cánh đồng theo ô vuông. Điểm gặt nọ cách điểm gặt kia đúng bằng khoảng cách đã tính.

Chọn trong máy móc, nếu các đơn vị của tổng thể chung được sắp xếp theo các tiêu thức không liên quan trực tiếp đến tiêu thức nghiên cứu thì sẽ có sai số ngẫu nhiên. Ngược lại, nếu các đơn vị của tổng thể chung được sắp xếp theo tiêu thức nghiên cứu thì lại xuất hiện sai số có hệ thống. Sai số có hệ thống lớn hay nhỏ phụ thuộc vào việc chọn đơn vị đầu tiên. Nếu đơn vị chọn đầu tiên ở gần giữa hoặc giữa khoảng cách đầu thì sai số hệ thống sẽ nhỏ. Vì lý do trên, người ta thường sắp xếp các đơn vị theo tiêu thức không có liên quan đến nội dung nghiên cứu.

Sai số bình quân chọn mẫu trong chọn máy móc cũng được tính như khi chọn ngẫu nhiên đơn thuần. Vì chọn máy móc là chọn một lần nên dùng công thức chọn một lần để tính.

Phương pháp chọn máy móc có thủ tục tiến hành đơn giản hơn chọn ngẫu nhiên đơn thuần. Do việc chọn theo từng khoảng cách nhất định nên số đơn vị mẫu được phân phối đều trong tổng thể chung, tính chất đại biểu của tổng thể mẫu được nâng cao. Trong nhiều trường hợp cụ thể người ta thấy sai số chọn mẫu của chọn máy móc thường nhỏ hơn so với sai số của chọn ngẫu nhiên đơn thuần.

8.2.2.3. Phương pháp chọn phân loại (chọn phân tổ)

Chọn phân loại là tiến hành chọn các đơn vị mẫu khi tổng thể chung đã được phân chia thành các tổ theo tiêu thức liên quan trực tiếp đến mục đích nghiên cứu. Việc chọn các đơn vị từ các tổ tiến hành theo phương pháp chọn ngẫu nhiên.

Chẳng hạn, khi điều tra đời sống cán bộ công nhân viên chức, trước hết phân chia cán bộ công nhân viên theo các ngành nghề khác nhau, sau đó từ trong các ngành chọn ra các hộ để điều tra.

Khi phân tổ, các tổ thường có quy mô khác nhau (rất ít trường hợp quy mô các tổ bằng nhau). Nếu chọn số đơn vị tổng thể mẫu trong mỗi tổ tương ứng với tỷ trọng của nó chiếm trong tổng thể thì gọi là chọn phân loại theo tỷ lệ. Ngược lại, nếu chọn số đơn vị mẫu trong các tổ đều bằng

nhau thì gọi là chọn không theo tỷ lệ. Trong chọn tỷ lệ, số đơn vị chọn trong mỗi tổ sẽ được tính theo công thức:

$$n_i = n \frac{N_i}{N}$$

Trong thực tế, người ta thường dùng phương pháp chọn phân loại theo tỷ lệ vì chọn như vậy sẽ được một tổng thể mẫu có kết cấu gần giống với kết cấu của tổng thể chung.

Sai số bình quân chọn mẫu trong chọn phân loại tỷ lệ không phụ thuộc vào phương sai chung mà phụ thuộc vào bình quân các phương sai tổ. Vì vậy, trong trường hợp chọn phân loại theo tỷ lệ ta có các công thức tính sau:

Nhiệm vụ suy rộng	Chọn nhiều lần	Chọn một lần
Suy rộng số trung bình:	$\mu_x = \sqrt{\frac{\sigma_1^2}{n}}$	$\mu_x = \sqrt{\frac{\sigma_1^2}{n} (1 - \frac{n}{N})}$
Suy rộng tỷ lệ:	$\mu_p = \sqrt{\frac{p(1-p)}{n}}$	$\mu_p = \sqrt{\frac{p(1-p)}{n} (1 - \frac{n}{N})}$

Trong các công thức trên:

$$\sigma^2 = \frac{\sum \sigma_i^2 N_i}{\sum N_i} = \sqrt{\frac{p(1-p)}{n}} = \overline{p(1-p)} = \frac{\sum p(1-p) N_i}{\sum N_i}$$

Thống kê toán đã chứng minh rằng số bình quân của các phương sai tổ nhỏ hơn hoặc bằng phương sai chung. Nghĩa là $\sigma_i^2 \leq \sigma^2$ và $\overline{p(1-p)} \leq p(1-p)$. Vì vậy, sai số bình quân chọn mẫu trong chọn phân loại nhỏ hơn trong chọn ngẫu nhiên đơn thuần và máy móc.

Trong trường hợp chọn phân loại không theo tỷ lệ, sai số bình quân chọn mẫu được tính theo công thức:

$$\mu = \frac{1}{N} \sqrt{\sum \mu_i^2 N_i}$$

- Trong đó: μ_i - Sai số bình quân chọn mẫu trong từng tổ.
 N_i - Số đơn vị trong từng tổ của tổng thể chung.

Muốn cho tổng thể mẫu có tính chất đại biểu cao hơn nữa, có thể dùng phương pháp rút mẫu tối ưu của Nây-man. Theo phương pháp rút mẫu tối ưu này, số đơn vị chọn ra ở mỗi tổ không những tương ứng với tỷ lệ của nó chiếm trong tổng thể mà còn tương ứng với độ biến thiên tiêu thức ở mỗi tổ. Nghĩa là số đơn vị mẫu trong mỗi tổ được tính theo công thức.

$$n_i = n \cdot \frac{\sigma_i N_i}{\sum \sigma_i N_i}$$

Trong chọn tối ưu, sai số bình quân chọn mẫu được tính theo công thức:

	Chọn nhiều lần	Chọn một lần
Suy rộng số trung bình	$\mu_x = \frac{1}{N} \frac{\sum \sigma_i N_i}{\sqrt{n}}$	$\mu_x = \frac{1}{N} \frac{\sum \sigma_i N_i}{\sqrt{n}} \sqrt{\left(1 - \frac{n}{N}\right)}$
Suy rộng tỷ lệ	$\mu_p = \frac{1}{N} \frac{\sum p(1-p) N_i}{\sqrt{n}}$	$\mu_p = \frac{1}{N} \frac{\sum p(1-p)}{\sqrt{n}} \sqrt{\left(1 - \frac{n}{N}\right)}$

8.2.2.4. Phương pháp chọn cả khối (mẫu chùm)

Trong các phương pháp tổ chức chọn mẫu trên, mỗi lần chọn chỉ rút ra từng đơn vị. Trong chọn cả khối mỗi lần chọn, số mẫu được rút ra không phải lẻ tẻ từng đơn vị mà từng nhóm (khối) đơn vị. Mỗi nhóm đơn vị chọn ra được điều tra hết không bỏ sót một đơn vị nào. Có thể nói đây là điều tra toàn bộ trong các khối được chọn ra.

Muốn chọn theo phương pháp này, trước hết tổng thể chung phải được chia thành các khối có quy mô bằng nhau hoặc khác nhau. Để tiện cho việc tính toán người ta thường chia thành các khối có quy mô bằng nhau. Việc chọn khối có thể tiến hành theo phương pháp chọn ngẫu nhiên đơn thuần hoặc chọn máy móc.

Trong chọn cả khối, sai số bình quân chọn mẫu không phụ thuộc vào phương sai chung mà phụ thuộc vào sự chênh lệch giữa các số trung bình khối hay còn gọi là phương sai các số trung bình khối. Vì vậy, sai số trung bình chọn mẫu trong chọn cả khối được tính như sau:

- Khi suy rộng số trung bình:

$$\mu_x = \sqrt{\frac{\sigma_x^2}{R-1} \left(\frac{R-r}{R-1} \right)}$$

Trong đó:

R: Số khối của tổng thể chung.

r: Số khối được chọn ra để điều tra.

Phương sai giữa các số trung bình khối được chọn, được tính bằng công thức:

$$\sigma_x^2 = \frac{\sum (x_i - x)^2 n_i}{\sum n_i}$$

Nếu số đơn vị các khối bằng nhau thì:

$$\sigma_x^2 = \frac{\sum (x_i - x)^2}{r}$$

Trong đó:

x_i : Là số trung bình của mỗi khối được chọn, x là số trung bình của tất cả các khối được chọn.

- Khi suy rộng tỷ lệ:

$$\mu_p = \sqrt{\frac{w_r(1-w_r)}{r} \left(\frac{R-r}{R-1} \right)}$$

Trong đó:

w_r : Tỷ lệ trung bình của các khối được chọn, được tính bằng công thức:

$$W_r = \frac{\sum W_i n_i}{\sum n_i}$$

Nếu số đơn vị các khối bằng nhau thì:

$$w_r = \frac{\sum w_i}{r}$$

Trong đó:

W_i : Là tỷ lệ từng khối được chọn.

Thống kê toán đã chứng minh rằng phương sai của các số trung bình khối luôn nhỏ hơn phương sai chung, cụ thể là:

$$\sigma_x^2 \leq \sigma^2$$

$$W_r (1 - W_r) \leq W (1 - W)$$

Vì vậy, sai số trung bình chọn mẫu trong chọn cả khối nhỏ hơn trong chọn ngẫu nhiên đơn thuần.

8.2.2.5. Phương pháp chọn mẫu phân tầng (chọn mẫu nhiều cấp hay chọn mẫu nhiều bậc)

Chọn mẫu phân tầng là phương pháp tổ chức chọn mẫu phải thông qua ít nhất hai cấp chọn trung gian. Đầu tiên cần xác định các đơn vị mẫu cấp I (có quy mô lớn), sau đó từ các tổ có quy mô lớn lại phân chia thành các đơn vị chọn mẫu cấp II (có quy mô nhỏ hơn) và cứ như thế cho đến cấp cuối cùng. Về bản chất, phương pháp này là sự biến thể của phương pháp chọn mẫu cả khối.

Ví dụ: Trong điều tra đời sống nhân dân, người ta thường tiến hành chọn theo ba bậc: Đầu tiên chọn 50% số tỉnh trong toàn quốc. Sau đó, trong các tỉnh đã chọn lấy ra 10% số huyện. Từ các huyện đã chọn, lấy ra 20% gia đình để điều tra. Như vậy, tỷ lệ chọn gia đình trong toàn quốc sẽ là:

$$p = p_1 \times p_2 \times p_3 = 0,5 \times 0,1 \times 0,2 = 0,01 \text{ hay } 1\%$$

Chọn phân tầng khác chọn phân loại ở chỗ, trong chọn phân loại người ta chọn đơn vị điều tra từ tất cả các tổ, còn trong chọn phân tầng chỉ chọn đơn vị ở những tổ đã chọn.

Sai số chọn mẫu trong trường hợp chọn từ ba bậc trở lên được tính theo công thức:

$$\mu^2 = \mu_1^2 + \frac{\mu_2^2}{n_1} + \frac{\mu_3^2}{n_1 \cdot n_2} + \frac{\mu_4^2}{n_1 \cdot n_2 \cdot n_3} + \dots$$

Trong đó:

μ_1, μ_2, μ_3 : Là sai số bình quân chọn mẫu ở từng bậc.

n_1, n_2, n_3 : Là số đơn vị mẫu ở từng bậc.

Ưu điểm cơ bản của chọn phân tầng là giảm bớt khối lượng công việc trong việc lập danh sách các đơn vị. Các đơn vị ở bậc đầu thường có quy mô lớn (tỉnh, huyện...) nên số đơn vị ít. Hơn nữa, các bậc tiếp sau chỉ cần lập danh sách những đơn vị đã được chọn ở bậc trước nên khối lượng công việc giảm đi.

Trong các phương pháp chọn mẫu ngẫu nhiên được trình bày ở trên thì phương pháp chọn mẫu phân loại, đặc biệt là phương pháp chọn tối ưu thường cho sai số là nhỏ nhất, song nó lại là phương pháp tổ chức chọn mẫu phức tạp nhất.

8.3. Quy trình một cuộc điều tra chọn mẫu ngẫu nhiên

Một cuộc điều tra chọn mẫu ngẫu nhiên thường được tiến hành qua các giai đoạn sau đây:

8.3.1. Xác định mục đích nghiên cứu

Xác định mục đích nghiên cứu tức là phải xác định một cách rõ ràng là cuộc điều tra đó nhằm tìm hiểu những vấn đề gì, phục vụ cho các yêu cầu cụ thể nào?

Xác định mục đích nghiên cứu là bước khởi đầu quan trọng, là tiền đề cho các giai đoạn sau.

8.3.2. Xác định tổng thể nghiên cứu

Tổng thể nghiên cứu là tổng thể chung bao gồm cả các đơn vị của hiện tượng nghiên cứu - tức xác định N. Để xác định tổng thể nghiên cứu phải dựa vào mục đích nghiên cứu khác nhau thì tổng thể nghiên cứu khác nhau. Thực chất của việc xác định tổng thể nghiên cứu là đi xác định các đơn vị của nó. Muốn vậy, ngoài việc dựa vào mục đích nghiên cứu còn phải dựa vào lý luận kinh tế - xã hội, tình hình thực tế để đưa ra định nghĩa, những tiêu chuẩn làm căn cứ để xem xét một cách cụ thể.

8.3.3. Xác định nội dung điều tra

Xác định nội dung điều tra là xác định danh mục các tiêu thức cần điều tra trên các đơn vị của tổng thể mẫu và được cụ thể hóa bằng phiếu (biểu) điều tra. Để xác định nội dung điều tra thì phải dựa vào mục đích

ngiên cứu. Mục đích nghiên cứu đòi hỏi phải giải quyết nhiều vấn đề thì nội dung điều tra phải bao gồm nhiều tiêu thức.

8.3.4. Xác định số lượng đơn vị của tổng thể mẫu và phương pháp tổ chức chọn mẫu

Xác định số lượng đơn vị của tổng thể mẫu tức là xác định n (còn gọi là kích thước mẫu, quy mô mẫu). Để xác định số lượng đơn vị của tổng thể mẫu cần phải cho trước phạm vi sai số chọn mẫu và xác suất suy rộng tài liệu (khi nghiên cứu các hiện tượng kinh tế - xã hội, thường lấy xác suất 0,9544 hay 95,44%). Các công thức tính số lượng đơn vị của tổng thể mẫu đã được trình bày ở trên.

Sau khi đã xác định được số lượng đơn vị của tổng thể mẫu, dựa vào đặc điểm của hiện tượng nghiên cứu và khả năng tổ chức điều tra để áp dụng các phương pháp tổ chức chọn mẫu đã được đề cập đến ở phần III nhằm xác định các đơn vị cụ thể của tổng thể mẫu (còn gọi là dàn chọn mẫu).

8.3.5. Tiến hành thu thập tài liệu ở các đơn vị của tổng thể mẫu

Dựa vào phiếu điều tra để tiến hành thu thập tài liệu ở các đơn vị của tổng thể mẫu. Có nhiều phương pháp thu thập tài liệu như: Phương pháp đăng ký trực tiếp, phương pháp phỏng vấn trực tiếp,... Tùy thuộc vào điều kiện và tính chất của cuộc điều tra để áp dụng phương pháp thu thập tài liệu cho phù hợp.

8.3.6. Suy rộng kết quả điều tra chọn mẫu

Sau khi đã thu thập được đầy đủ tài liệu ở các đơn vị của tổng thể mẫu, căn cứ vào đó để tiến hành tính toán và suy rộng (còn gọi là ước lượng) các đặc điểm của tổng thể chung. Có hai phương pháp suy rộng là: Suy rộng trực tiếp và suy rộng khoảng.

Suy rộng trực tiếp là coi các mức độ của tổng thể mẫu cũng là các mức độ của tổng thể chung. Ví dụ: Coi bình quân mẫu, tỷ lệ của tổng thể mẫu cũng là bình quân, tỷ lệ của tổng thể chung.

Suy rộng khoảng là các mức độ của tổng thể chung được xác định nhận giá trị trong một khoảng nào đó với xác suất cho trước.

8.3.7. Đưa ra kết luận về tổng thể chung

Đây là giai đoạn cuối cùng thể hiện kết quả của quá trình nghiên cứu. Câu hỏi được đặt ra là: “Có thể có các kết luận gì về tổng thể chung?”, “Các kết luận đó có đáp ứng được mục đích nghiên cứu đã được đặt ra hay không? Có phản ánh được đặc điểm, bản chất của hiện tượng hay không?”. Từ đó cần đưa ra những giải pháp, kiến nghị cụ thể để thúc đẩy sự phát triển của hiện tượng.

Các kết quả nghiên cứu đạt được cần được trình bày thông qua các bảng thống kê, các đồ thị thống kê và báo cáo phân tích.

8.4. Điều tra chọn mẫu phi ngẫu nhiên

Điều tra chọn mẫu phi ngẫu nhiên là điều tra chọn mẫu mà trong đó các đơn vị của tổng thể mẫu được chọn ra trên cơ sở phân tích đặc điểm của hiện tượng và kinh nghiệm thực tế. Do đó, để đảm bảo chất lượng của tài liệu điều tra, cần phải giải quyết tốt các vấn đề sau đây:

8.4.1. Phân tổ chính xác hiện tượng nghiên cứu

Hiện tượng nghiên cứu thường có kết cấu phức tạp, gồm nhiều tổ, nhiều bộ phận có đặc điểm và tính chất khác nhau. Trên cơ sở phân tổ chính xác hiện tượng nghiên cứu, các đơn vị có đặc điểm và tính chất giống nhau (hoặc gần giống nhau) sẽ được đưa vào một tổ. Từ mỗi tổ sẽ chọn ra các đơn vị đại diện (còn gọi là điển hình) cho tổ đó. Tập hợp các đơn vị đại diện của các tổ tạo thành tổng thể mẫu.

Giả sử điều tra về mức sống ở một địa phương, có thể phân mức sống thành ba tổ là nghèo, trung bình, khá giả. Từ mỗi tổ sẽ chọn ra một số hộ đại diện cho tổ đó để tiến hành điều tra.

8.4.2. Xác định số lượng đơn vị cần điều tra

Việc xác định số lượng đơn vị cần điều tra phải căn cứ vào tính chất phức tạp của hiện tượng nghiên cứu, lực lượng cán bộ, kinh phí cho cuộc điều tra, ... Từ đó, quyết định điều tra bao nhiêu đơn vị, số lượng các đơn vị được chọn ra để điều tra phải đủ lớn, có thể đại diện cho toàn bộ hiện tượng.

8.4.3. Lựa chọn các đơn vị điều tra

Các đơn vị được lựa chọn để điều tra thực tế thường là những đơn vị có mức độ xấp xỉ với mức độ bình quân của tổ. Khi lựa chọn các đơn vị để điều tra thực tế cần phải thông qua việc phân tích, bàn bạc tập thể của những người có kinh nghiệm, am hiểu tình hình thực tế.

8.4.4. Suy rộng kết quả điều tra

Sau khi đã thu thập được tài liệu ở các đơn vị điều tra thì tiến hành tính toán suy rộng trực tiếp cho toàn bộ hiện tượng. Vì các đơn vị điều tra được lựa chọn đại diện cho từng tổ nên khi suy rộng phải chú ý đến tỷ trọng của mỗi tổ chiếm trong toàn bộ hiện tượng.

TÓM TẮT CHƯƠNG 8

Điều tra chọn mẫu là một loại điều tra không toàn bộ, trong đó người ta chỉ chọn ra một số đơn vị trong toàn bộ các đơn vị của đối tượng nghiên cứu để điều tra thực tế. Có 2 phương pháp điều tra chọn mẫu: điều tra chọn mẫu ngẫu nhiên và điều tra chọn mẫu phi ngẫu nhiên. Điều tra chọn mẫu cho phép tiết kiệm chi phí, đảm bảo tính kịp thời của tài liệu điều tra.

Ta sử dụng các công thức toán học để xác định sai số trung bình chọn mẫu, phạm vi sai số chọn mẫu, số đơn vị tổng thể mẫu. Có hai phương pháp suy rộng tài liệu: Phương pháp tính đổi trực tiếp và phương pháp hệ số điều chỉnh. Tùy thuộc vào đặc điểm của tổng thể chung và khả năng tổ chức điều tra, có thể sử dụng các phương pháp để lựa chọn đơn vị tổng thể mẫu sau đây: Chọn ngẫu nhiên, chọn đơn thuần, chọn máy móc, chọn phân loại, chọn cả khối. Phạm vi sai số chọn mẫu đối với từng phương pháp có sự khác nhau.

Một cuộc điều tra chọn mẫu thường được tiến hành theo một quy trình gồm các giai đoạn sau: Xác định mục đích của cuộc điều tra chọn mẫu, xác định tổng thể nghiên cứu, xác định nội dung điều tra, xác định số lượng, đơn vị tổng thể mẫu và phương pháp tổ chức chọn mẫu tiến hành thu thập tài liệu trên mẫu, suy rộng kết quả điều tra chọn mẫu và đưa ra kết luận về tổng thể chung.

CÂU HỎI ÔN TẬP CHƯƠNG 8

1. Trình bày khái niệm, ý nghĩa và trường hợp vận dụng của điều tra chọn mẫu?
2. Trình bày khái niệm, các nhân tố ảnh hưởng đến sai số chọn mẫu, công thức xác định sai số trung bình chọn mẫu?
3. Trình bày phương pháp xác định phạm vi sai số chọn mẫu, quy mô tổng thể mẫu và các phương pháp suy rộng kết quả điều tra chọn mẫu?
4. Trình bày các phương pháp tổ chức chọn mẫu thường dùng, trường hợp vận dụng của các hình thức tổ chức đó?
5. Trình bày quy trình tổ chức điều tra chọn mẫu. Cho ví dụ minh họa?

BÀI TẬP CHƯƠNG 8

Bài 1:

Trong một doanh nghiệp có 1000 công nhân, người ta chọn ra 50 công nhân để điều tra năng suất lao động theo phương pháp chọn 1 lần. Kết quả điều tra cho thấy: Thời gian hao phí trung bình để sản xuất một sản phẩm của mỗi công nhân là 32 phút, độ lệch tiêu chuẩn là 6 phút.

Hãy tính:

Phạm vi sai số chọn mẫu khi suy rộng kết quả điều tra, với trình độ tin cậy 0,954?

Thời gian hao phí trung bình để sản xuất 1 sản phẩm của mỗi công nhân trong doanh nghiệp?

Với số liệu trên, nếu điều tra 100 công nhân thì kết quả suy rộng sẽ là bao nhiêu. Hãy so sánh các kết quả tính được?

Bài 2:

Một doanh nghiệp có 2000 công nhân, người ta chọn 100 công nhân để điều tra về năng suất lao động theo phương pháp chọn ngẫu nhiên đơn thuần (không hoàn lại). Kết quả điều tra như sau:

NSLĐ (Tấn)	Số công nhân (Người)
40 - 50	15
50 - 60	60
60 - 70	25

Hãy tính:

- Năng suất lao động trung bình của số công nhân đã được điều tra?
- Phương sai mẫu về năng suất lao động?
- Phạm vi sai số chọn mẫu khi suy rộng NSLĐ trung bình của công nhân trong toàn doanh nghiệp với độ tin cậy 0,954?
- Nếu số công nhân được chọn theo phương pháp chọn ngẫu nhiên đơn thuần (có hoàn lại) thì sai số trung bình chọn mẫu sẽ là bao nhiêu?

Bài 3:

Theo kết quả điều tra 10% nhân khẩu của thành phố A vào tháng 1 năm N bằng phương pháp chọn ngẫu nhiên đơn thuần (không hoàn lại) thấy có 8% số nhân khẩu trên 80 tuổi; 14% dưới 6 tuổi và 18% là công nhân.

Biết rằng tổng số nhân khẩu của thành phố A tại thời điểm điều tra là 300.000 người.

Với xác suất tin cậy 0,954 hãy tính:

Phạm vi sai số chọn mẫu về tỷ trọng của mỗi loại nhân khẩu nói trên?

Tỷ trọng của mỗi loại nhân khẩu nói trên trong cả thành phố?

Bài 4:

Một doanh nghiệp sản xuất bóng đèn điện, người ta tiến hành chọn mẫu 200 bóng đèn để xác định thời gian thấp sáng trung bình một bóng đèn, trong một đợt sản xuất gồm 100.000 bóng theo phương pháp chọn không hoàn lại. Với xác suất 0,997, hãy tính:

Phạm vi sai số chọn mẫu về thời gian thấp sáng trung bình của bóng đèn? Biết độ chênh lệch tiêu chuẩn của số bóng đèn điều tra là 100 giờ.

Thời gian thấp sáng trung bình của số bóng đèn đã sản xuất? Nếu thời gian thấp sáng trung bình của số bóng đèn điều tra là 980 giờ.

TÀI LIỆU THAM KHẢO

1. Nguyễn Hữu Hòe. *Giáo trình Lý thuyết Thống kê*. Nhà xuất bản Thống kê - 1984.
2. Kim Ngọc Huynh. *Bài giảng và bài tập Lý thuyết Thống kê*. Nhà xuất bản Đại học và Giáo dục chuyên nghiệp - 1992.
3. Tô Thị Phượng. *Lý thuyết Thống kê*. Nhà xuất bản Thống kê - 1998.
4. Trần Thị Kim Thu. *Lý thuyết Thống kê*. Nhà xuất bản Thống kê - 2006.
5. Viện Khoa học Thống kê. *Một số vấn đề phương pháp luận thống kê*. Nhà xuất bản Thống kê - 2005.
6. Tăng Văn Khiên. *Lý thuyết và ứng dụng*. Nhà xuất bản Thống kê - 2015.
7. *Luật Thống kê và các văn bản hướng dẫn thi hành*. Nhà xuất bản Thống kê - 2004.
8. Trần Thị Kim Thu. *Giáo trình Lý thuyết Thống kê*. Nhà xuất bản Đại học Kinh tế quốc dân - 2012.
9. P. A. Smoilova. *Lý thuyết thống kê*. Nhà xuất bản Thống kê Matxcova - 2008.
10. B.E Glinxki. *Phân tích Thống kê*. Nhà xuất bản "Flin" - 1998.

GIÁO TRÌNH

NGUYÊN LÝ THỐNG KÊ

Chịu trách nhiệm xuất bản:

GIÁM ĐỐC - TỔNG BIÊN TẬP
ThS. ĐỖ VĂN CHIẾN

Biên tập:

VƯƠNG NGỌC LAM

Trình bày:

TRẦN MẠNH HÀ - BÙI DŨNG THẮNG

Sách được xuất bản tại:

NHÀ XUẤT BẢN THỐNG KÊ

Địa chỉ: 86 Thụy Khuê, Tây Hồ, Hà Nội

Tel: (04) 38 471 483; Fax: (04) 38 474 185

Website: nxbthongke.com.vn

E-mail: nxbtk@gso.gov.vn; xuatbanthongke@gmail.com

In 1.000 cuốn, khổ 16 × 24cm, tại Nhà xuất bản Thống kê - Công ty Cổ phần In Khoa học công nghệ mới
Địa chỉ: 181 Lạc Long Quân, Nghĩa Đô, Cầu Giấy, TP. Hà Nội.
Số xác nhận ĐKXB: 1149-2016/CXBIPH/06-10/TK do Cục Xuất bản - In và phát hành
cấp ngày 22/4/2016. QĐXB số 56/QĐ-NXBTK ngày 27/5/2016
của Giám đốc - Tổng biên tập NXB Thống kê.
In xong, nộp lưu chiểu: tháng 8 năm 2016.